DOCUMENT RESUME

ED 315 920 EA 021 709

AUTHOR Oakes, Jeannie
TITLE Educational Indicators: A Guide for Policymakers. CPRE Occasional Paper Series.
INSTITUTION Center for Policy Research in Education.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
REPORT NO CPRE-OPE-01
PUB DATE Oct 86
GRANT OERI-6-86-0011
NOTE 49p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Trends; Elementary Secondary Education; *Outcomes of Education; *Politics of Education; *Program Implementation; *Statistical Analysis
IDENTIFIERS *Indicators

ABSTRACT

An educational indicator is a statistic revealing something about the education system's health or performance. Indicators must meet certain substantive and technical standards that define the kind of information they should provide and the features they should measure. There are two types of statistical indicators. Whereas single statistics provide readings about education (such as class size or number of schools using microcomputers), composite statistics (such as the pupil/teacher ratio) provide information concerning relationships among factors. Indicator systems measure separate components and interactions between components. Complete educational indicator systems are nonexistent. Indicators can supposedly help to: (1) report the status of American schooling and make district, state, and international comparisons; (2) monitor changes over time; (3) explain the causes of various conditions and changes; (4) predict likely changes; (5) profile system strengths and weaknessses; and (6) suggest improvement strategies for policymakers. Some of these claims can be met, while others are unrealistic, as this report shows. Indicator data are unlikely to produce unequivocably good or bad news and will be open to various interpretations. Decisions (mostly political) about desired schooling outcomes and conditions will determine the nature of any indicator system. There are numerous implementation issues, including desired level of information, the need to make fair comparisons, appropriate scope, political pressures, and reconciliation concerning priorities among policymakers and professional educators. A list of information sources is provided. (MLH)

Educational Indicators
A GUIDE FOR POLICYMAKERS

Jeannie Oakes

CPRE CENTER FOR POLICY RESEARCH IN EDUCATION

CENTER FOR POLICY RESEARCH IN EDUCATION

The Center for Policy Research in Education (CPRE), sponsored by the U.S. Department of Education, unites the Eagleton Institute of Politics at Rutgers University, The RAND Corporation, and the University of Wisconsin-Madison in a joint effort to improve the quality of schooling through research on state and local education policy.

CPRE has the following goals:

- To strengthen the connections among policy, practice, and performance in American elementary and secondary education,
- To focus education policy research on the central concerns of schooling—teaching, learning, school organization, and student performance,
- To produce knowledge useful to policymakers and their constituents,
- To broaden the range of options from which education policymakers can choose,
- To foster a dialogue between the consumers and producers of education policy research.

To ensure that the Center's work is useful to policymakers, CPRE's research agenda is organized about three questions fundamental to any assessment of education policy.

1. Will the policy improve the quality of schools and student performance?
2. Can it be implemented within given resource constraints?
3. Will it be supported in its implementation?

CPRE Occasional Paper Series

Occasional Papers are issued by CPRE to facilitate the exchange of ideas among policymakers and researchers who share an interest in education policy. The views expressed in these papers are those of individual authors, and are not necessarily shared by the U.S. Department of Education, The RAND Corporation, Rutgers University, or the University of Wisconsin-Madison. This publication was funded by the U.S. Department of Education, Office of Educational Research and Improvement, grant number OERI-6-86-0011.

Educational Indicators

A GUIDE FOR POLICYMAKERS

Jeannie Oakes

OPE-01

October 1986

CPRE CENTER FOR POLICY RESEARCH IN EDUCATION
RUTGERS UNIVERSITY • THE RAND CORPORATION • UNIVERSITY OF WISCONSIN-MADISON

ACKNOWLEDGMENTS

A number of people have contributed to the preparation of this brochure. Lorraine McDonnell of The RAND Corporation and Susan Fuhrman of Rutgers University originated the idea of an indicator handbook as a publication of the Center for Policy Research in Education and offered valuable ideas along the way. Helpful comments were also generously provided by Stephen Kaagan, Vermont Commissioner of Education; Representative Annette Morgan, Missouri House of Representatives; Ramsay Selden, Council of Chief State School Officers; Richard Shavelson, UCLA and RAND; and Marshall Smith, Stanford University. Mary Vaiana supplied considerable assistance in organizing and presenting the text. The author, of course, remains responsible for any factual or interpretive errors.

EDUCATIONAL INDICATORS: A GUIDE FOR POLICYMAKERS

Dropout rates exceed 60% in many major urban high schools.

Less than 10% of the top-scoring students on the SAT choose teaching as a career; most of them leave teaching within the first 7 years.

One out of every 4 children who began formal schooling in 1986 comes from a family living below the poverty line.

Policymakers confront such compelling statistics every day and often make decisions based on them. Both the confrontation and the decisions generate a host of new concerns.

- How should these numbers be interpreted—for example, how are dropout rates being calculated?
- What information from national statistics can be applied to our state or local district?
- What data are available for judging the quality of our educational system? If data do not exist, do we know how to obtain them?
- Can statistics be used to determine whether a policy is having its intended effect?
- What data about our schools should we collect on an ongoing basis?

In the past few years years, the annual publication of the Secretary of Education's "Wall Chart" has heightened these concerns about the appropriate uses of education statistics. Billed as a collection of *educational indicators*, the "Wall Chart" presents a series of statistics that compares states on a number of dimensions—e.g., students' SAT and ACT scores, graduation rates, teachers' salaries, pupil/teacher ratios, expenditures, and student population characteristics. The chart has been greeted with charges of unfair comparisons, inappropriate measures, and too little information to accurately portray or compare education. Nevertheless, each year the "Wall Chart" statistics make headlines, and policymakers have been pressured to respond. Not surprisingly, one important response has been the launching of intensive efforts to develop a better set of indicators to report the condition of the nation's schooling systems.

In the wake of the "Wall Chart" controversy, policymakers need to understand the uses and abuses of the education system's principal diagnostic tools—educational indicators. Those in positions to use indicators and pay for their collection must understand what the realistic information returns are from investing in them. This brochure is designed to help education decisionmakers understand the legitimate roles indicators might play in monitoring the condition of the educational system, tracking changes over time, and anticipating future change. It defines educational indicators, explains their principal applications, describes some of their limitations, and reviews the current state of the indicator art.

SUMMARY

For the convenience of the reader, we provide the following brief summary and guide to the more detailed discussion in the body of this brochure.

WHAT IS AN EDUCATIONAL INDICATOR?

An educational indicator is a statistic that tells something about the performance or health of the education system. For a statistic to be an indicator, it must have a standard against which it can be judged. Indicators must meet certain substantive and technical standards that define the kind of information they should provide and the features they should measure.

For details, see pages 1–2

TYPES OF INDICATORS

There are two types of indicators: single statistics and composite statistics. Both can be used individually or incorporated into indicator systems.

Single statistics provide readings about the educational system. Examples include class size or number of schools using microcomputers. Composite statistics, such as pupil/teacher ratio, provide information about relationships among factors. Indicator systems measure distinct components of the system, but also provide information about how those components interact. Complete educational indicator systems do not yet exist—in large part because no single model of the education process has gained widespread acceptance. The closest existing approximation of a national indicator system is the Secretary of Education's controversial "Wall Chart," consisting of a small number of statistics drawn from federal or aggregated state data. However, despite the lack of a comprehensive system model, far better indicators can be constructed that reflect widely shared views about essential schooling outcomes (such as higher student achievement) and important schooling conditions (such as well-trained and experienced teachers.)

For details, see pages 3–11

HOW INDICATORS ARE USED

Claims have been made that indicators will (1) report the status of American schooling, contrast schooling in different states and districts, and compare U.S. schooling with systems abroad; (2) monitor changes over time; (3) explain the causes of various conditions and changes; (4) predict likely changes in the future; (5) profile the strengths and weaknesses of the system; and (6) inform policymakers about how to improve the system.

Some of these claims can be met, others are unrealistic. And in all cases, the claims must be viewed with caution. For example:

- Indicators such as average class size and achievement test scores are now used to compare schooling across schools, states, and

regions; however, these comparisons are inappropriate unless the indicators are defined and measured in the same way.

- Indicators can track schooling trends if they measure enduring features of the system and are consistently defined; however, indicators will not, by themselves, tell us what caused the trend.
- When research has clearly linked indicators and other events in the education system, indicators can predict future performance quite accurately. More often, indicators may reveal patterns of change that allow us to predict with somewhat less certainty.
- A national indicator system could provide information about generic types of policies; however, state and local policymakers will need indicators focused on their specific conditions and needs.

For details, see pages *12-23*

INDICATORS IN A POLICY CONTEXT

Decisions about desired schooling outcomes and conditions will determine the nature of any indicator system. In large part, these decisions will be political. Implementation issues include the following:

- *Level of Information*: Local, state, and federal policymakers have different information needs, and each level will want an indicator system to focus on its policies. A basic decision is whether a national indicator system should be "top-down" or "bottom-up."
- *Making Fair Comparisons*: Not all states, schools, teachers, and students start out even. An indicator system must reflect these differences without institutionalizing lower expectations for some schools and their students.
- *Scope of an Indicator System*: Comprehensive indicator systems at either the federal or state level would be very costly and would impose substantial burdens on those supplying data. Very extensive data might well take so long to collect and analyze that the information would lose its timeliness. In addition, there is a trade-off between comprehensiveness of the system and frequency of data collection. Thus indicator systems will probably focus on only a few key measures.
- *Political Pressures*: The existence of an indicator system will impose political pressures. The education community will exert its own pressures to influence the measures by which it will be judged. Educators may feel pressure to bias responses in ways beneficial to their schools. The information that an indicator system provides could influence the public's willingness to support education initiatives or to push for implementing particular policies.
- *Who Should Make Decisions about Indicators*: Definitions of indicators follow from views about appropriate educational goals and the ways to implement them. Many different views can be found within the education community. Thus arriving at precise definitions of indicators will require reconciling potential conflicts among the priorities of policymakers and professional educators, and among federal, state, and local levels.

For details, see pages *24-32*

STATE OF THE INDICATOR ART

Substantial amounts of educational data are already collected regularly at all levels. The data vary in quality and in regularity of collection.

Researchers and practitioners are working to develop a better set of national educational indicators. In particular, they are striving to develop better educational measures and to identify existing measures that meet "good indicator" criteria. The most likely system to emerge from their efforts is an overlapping two-level system. National indicators would monitor core features of schooling across states; separate indicator systems within each state would collect information tailored to its goals and needs.

Indicators will bring considerable new knowledge to bear on educational issues, stimulate debate, and suggest new approaches. But they will not provide unequivocal judgments about the quality of schooling or direct answers about how education should be improved.

For details, see pages 33–37

WHAT IS AN EDUCATIONAL INDICATOR?

SOME DEFINITIONS

An *educational indicator* is a statistic about the educational system that reveals something about its performance or health. Like the odometer, speedometer, temperature, and fuel gauges in a car, educational indicators provide essential information about the system's current functioning, suggest whether good progress is being made, and warn of potential problems. Obviously, indicators do not tell everything about a system. Instead, they provide an "at a glance" indication of current conditions and may even augur future prospects.

Not all statistics are indicators. To provide information about the health of a system, the indicator must have a reference point, some other measure or standard against which to judge the statistic. In some cases, setting the reference points is straightforward. For example, the car's temperature gauge has readings that represent a normal operating range for the engine. However, in a social system like education, setting the reference point is a more subjective undertaking. Usually, the reference point is some socially agreed-upon standard, a past reading, or a comparison with another location. Of course, comparing readings across time or locations doesn't tell us whether conditions are bad or good; it only tells us whether things are better or worse.

CHARACTERISTICS OF INDICATORS

Educational indicators are designed to provide insight into the health and effectiveness of the system, and to be useful in the policy context. To meet these goals, a good educational indicator should satisfy the following substantive and technical criteria.

Substantive Criteria

Indicators should provide *at least one* of the following types of information:

- Information that describes the educational system's *performance* in achieving desired educational conditions and outcomes—for example, its progress toward higher achievement test scores, decreased dropout rates, greater equality of opportunity, and safer school environments. Indicators of this type provide *benchmarks* for measuring progress.
- Information about *features of the system known to be linked with desired outcomes*—for example, resources such as instructional time that research has identified as relating to students' achievement. Indicators of this type can be seen as leading indicators. They have predictive value because when they change, other changes can be expected to follow.
- Information that describes *central features of the system*—for example, the amount of financial resources available, teachers' work load, and schools' curriculum offerings. Even though research has not as yet determined the relationship of these features to particu-

lar outcomes, we need information about these features to better understand how the system works.

- Information that is *problem-oriented*. Indicators are needed that provide information about current or potential problems in the system—for example, factors linked to teacher supply and demand, or to the special circumstances of schools in poor urban centers.
- Information that is *policy relevant*. Indicators should describe educational conditions of particular concern to policymakers and amenable to change by policy decisions. For example, indicators of teacher characteristics such as educational background and training would be policy-relevant because they can be changed through legislation or regulations governing teacher licensing. Most policy-relevant indicators will satisfy at least one of the other substantive criteria as well.

Technical Criteria

In addition to the substantive criteria, indicators should have the following technical characteristics.

- Indicators should measure *ubiquitous* features of schooling—those dimensions that can be found in some form throughout the system. That way, information can be collected and compared across diverse settings such as school types and locales. For example, statistics such as per-pupil expenditures or pupil/teacher ratios could be used as measures of fiscal and human resources available across schools and districts of various levels and types, across states, and even internationally.
- Indicators should measure *enduring* features of the system—for example, the number of courses in mathematics required for high school graduation—rather than specific policies such as whether computer math is a graduation requirement. If more generic measures are used, important features can be traced over time and trends analyzed.
- Indicators should be *readily understood* by a broad audience of educators and policymakers.
- Indicators should be *feasible* in terms of time, costs, and expertise required to collect the relevant information.
- Indicators should be generally accepted as *valid and reliable* statistics. They should measure what they are intended to measure, and they should do so consistently. *Generally accepted* are important words here because, as we will see in the next section, the choice of an indicator is determined by one's model—implicit or explicit—of how the education system works. To date, no single model has gained universal acceptance. So while researchers struggle to improve our understanding of the educational system, policymakers and practitioners must proceed by using indicators of those features we do understand and others they believe to be important.

TYPES OF INDICATORS

There are two basic types of indicators: single statistics and composite statistics. Both types can be used individually, and they may be incorporated into indicator systems. In this section, we describe and illustrate both types, illustrate how they might fit together in an indicator system, and discuss in more detail the relationship between indicators and models of the education system.

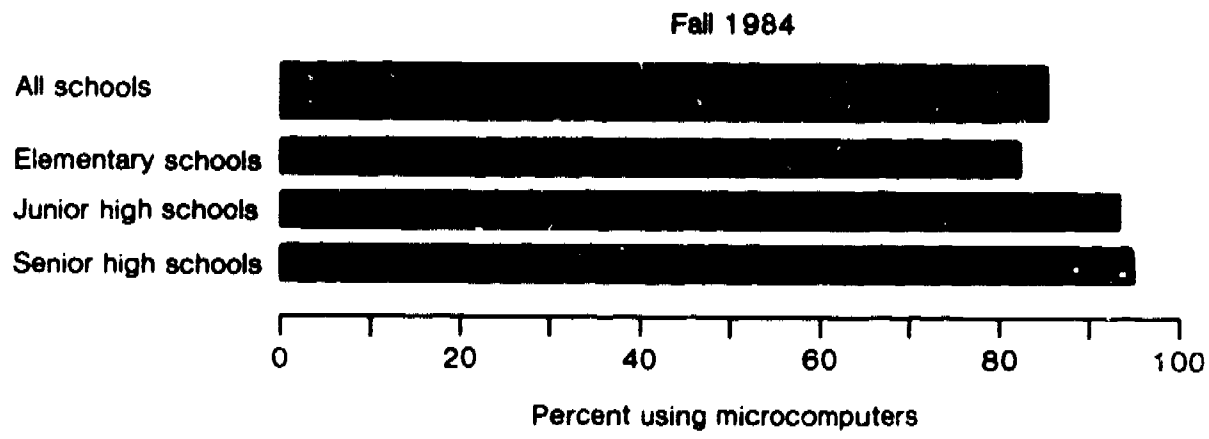
SINGLE STATISTICS AS INDICATORS

Single statistics can often be useful indicators. For example, we judge health by body temperature, cholesterol level, or weight; fitness from heart rate or percentage of body fat; athletic prowess by points scored or speed.

Educational indicators might include such single statistics as total education expenditures, class size, the average salary earned by a teacher with a Master's degree and 5 years of teaching experience or even the extent to which instructional aids such as computers are being used (shown in Fig. 1). Each of these single statistics might qualify as an indicator of the level of resource commitment to education.

COMPOSITE STATISTICS AS INDICATORS

An indicator may also be a composite statistic that measures and reports combinations of related events or characteristics. For example, to assess the economy, the Bureau of Labor Statistics develops several composite indicators. The Gross National Product is a total of the value of all the goods and services produced. Combined, these data form an index that provides information about the general level of productivity in the nation. Similarly, to estimate the cost of living, the Bureau compiles the Consumer Price Index by combining food prices, rents, and the costs of



Source: NCES, *The Condition of Education*, 1985.

Fig. 1—Single statistic indicator

a number of other goods and services. Unlike the GNP, however, the CPI is computed as a *weighted* index of a sample of price changes. The weights enable the scale to reflect accurately the magnitude or importance of price effects in various sectors of the economy—for example, the housing market or energy prices. Appropriately weighted, the combination of prices provides an estimate of living costs and tracks changes in those costs over time.

Composite statistical indices can also be useful as educational indicators. For example, the Department of Education has constructed a composite index of states' "educational need" by combining key measures of students' socioeconomic characteristics (shown in Fig. 2). Many policymakers also want an indicator that reflects the qualifications of the teaching force, but no single statistic can provide sufficient information. Therefore, we might construct a composite statistic of teacher qualifications by combining a number of separate qualification dimensions into a scale. We might add together the number and level of earned degrees, years of teaching experience, and number of hours spent in post-certification coursework and inservice training. If it seemed appropriate, we could weight some factors—for example, a Master's degree might be counted more heavily than participation in inservice training. If constructed well, a score on the Teacher Qualification Index would provide a more accurate picture of teachers' qualifications than could be obtained with single measures.

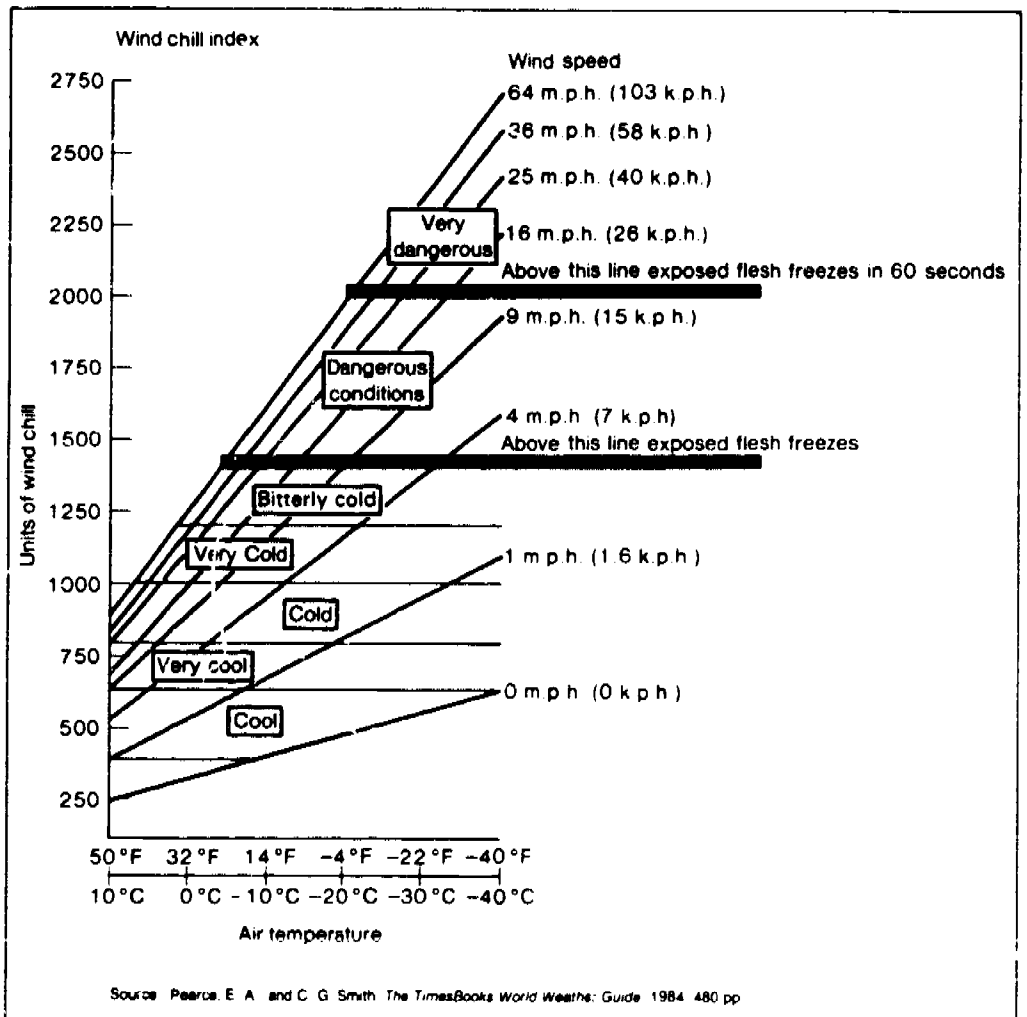
	Composite Index of Educational Service Requirements				
State and Region	Percent Children 5-17 in Poverty 1980	Percent Handicapped Children 1984	Percent Limited-English Proficient Children 1980	Index of Educational Service Requirements	Classification on Index
United States	16.3	10.9	9.6		
New England					
Connecticut	10.4	13.7	3.1	0.3	Low
Maine	16.1	13.9	3.1	11.0	Moderate
Massachusetts	12.3	13.9	3.0	11.0	Moderate
New Hampshire	8.9	9.6	3.1	6.0	Low
Rhode Island	12.6	13.5	4.3	11.5	Moderate
Vermont	13.0	10.9	2.2	10.0	Moderate
Mideast					
Delaware	14.6	16.4	2.4	11.0	Moderate
District of Columbia	36.3	8.1	2.3	12.0	High
Maryland	11.9	13.3	2.2	8.0	Low
New Jersey	13.3	14.4	6.3	11.5	Moderate
New York	17.9	10.5	14.3	14.0	High
Pennsylvania	13.2	11.3	3.1	10.0	Moderate
Great Lakes					
Illinois	14.1	13.9	3.9	11.0	Moderate
Indiana	11.0	10.3	2.2	7.0	Low
Michigan	12.4	9.1	1.4	9.0	Moderate
Ohio	12.2	11.0	1.9	10.0	Moderate
Wisconsin	9.6	9.3	0.9	6.0	Low

Source: U.S. Department of Education, NCES, *Indicators of Education Statistics and Trends*, 1986.

Fig. 2—Composite statistic indicator

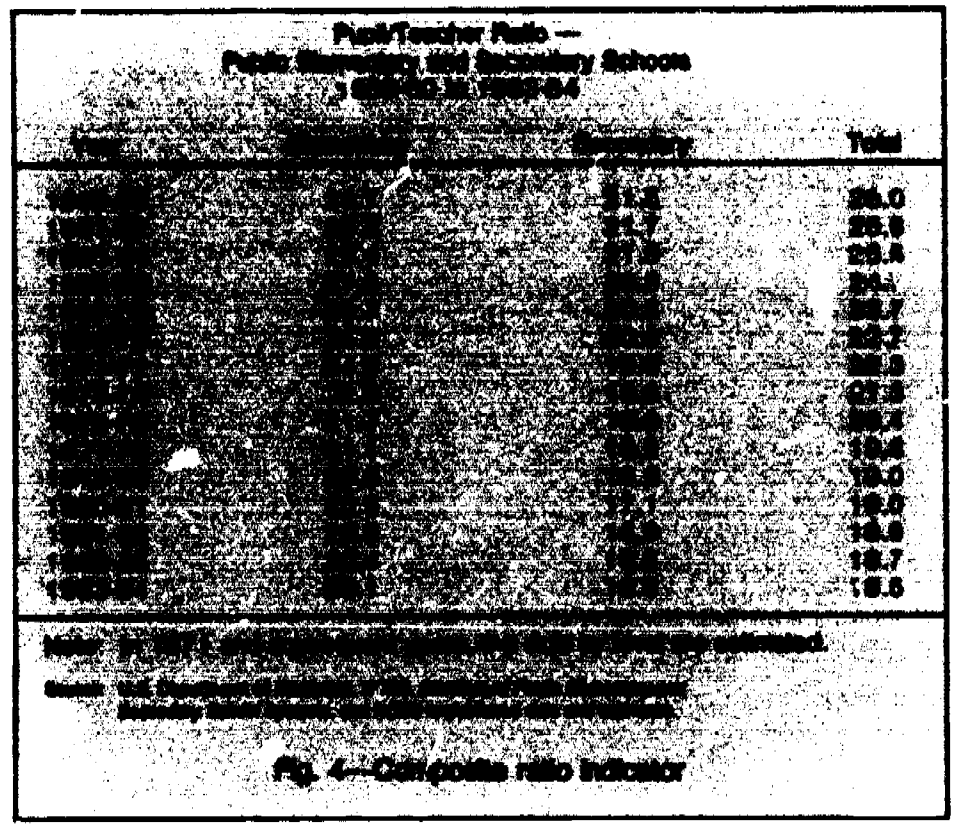
Composite indicators can also be constructed to provide information about *relationships* between two or more varying factors. A familiar example from weather reporting is the wind chill factor (see Fig. 3). The wind chill factor is a ratio that reports the condition—coldness—that results from the relationship between temperature and wind speed. Knowing wind speed in addition to temperature is important: The faster the wind blows, the faster the body loses heat. But this effect cannot be accurately estimated by simply adding the two numbers. Therefore, the wind chill factor captures the relationship of these two conditions and is a far better indicator of coldness than either of the two conditions separately.

Composite statistics that assess relationships seem particularly appropriate in education, where much of what happens results from several circumstances and events taking place at the same time. A common and quite simple example of such an indicator is pupil/teacher ratio (Fig. 4). Because it reveals the relationship between the number of students and the number of teachers in a particular state, district, or school, it



Source: Pearce, E. A. and C. G. Smith. *The TimesBooks World Weather Guide*. 1984. 480 pp.

Fig. 3—A composite indicator



provides more relevant information about the condition of education than if we simply considered the numbers of students and teachers separately. However, even this simple composite indicator presents difficulties since states often define "teachers" differently when calculating these ratios.

Although it would be more complicated than calculating pupil/teacher ratios, we might also want to construct a composite indicator to assess the academic performance of senior high schools since academic performance can not be accurately judged by one measure such as achievement test scores. School performance is a complex condition that includes much more than how well students score on tests. An accurate indicator of high schools' performance would probably require assessing the relationship of several factors—the scores of the 12th grade class when they entered the school as 9th or 10th graders, the percentage of dropouts between 9th and 12th grade, and the scores of 12th graders, to name just a few important components.

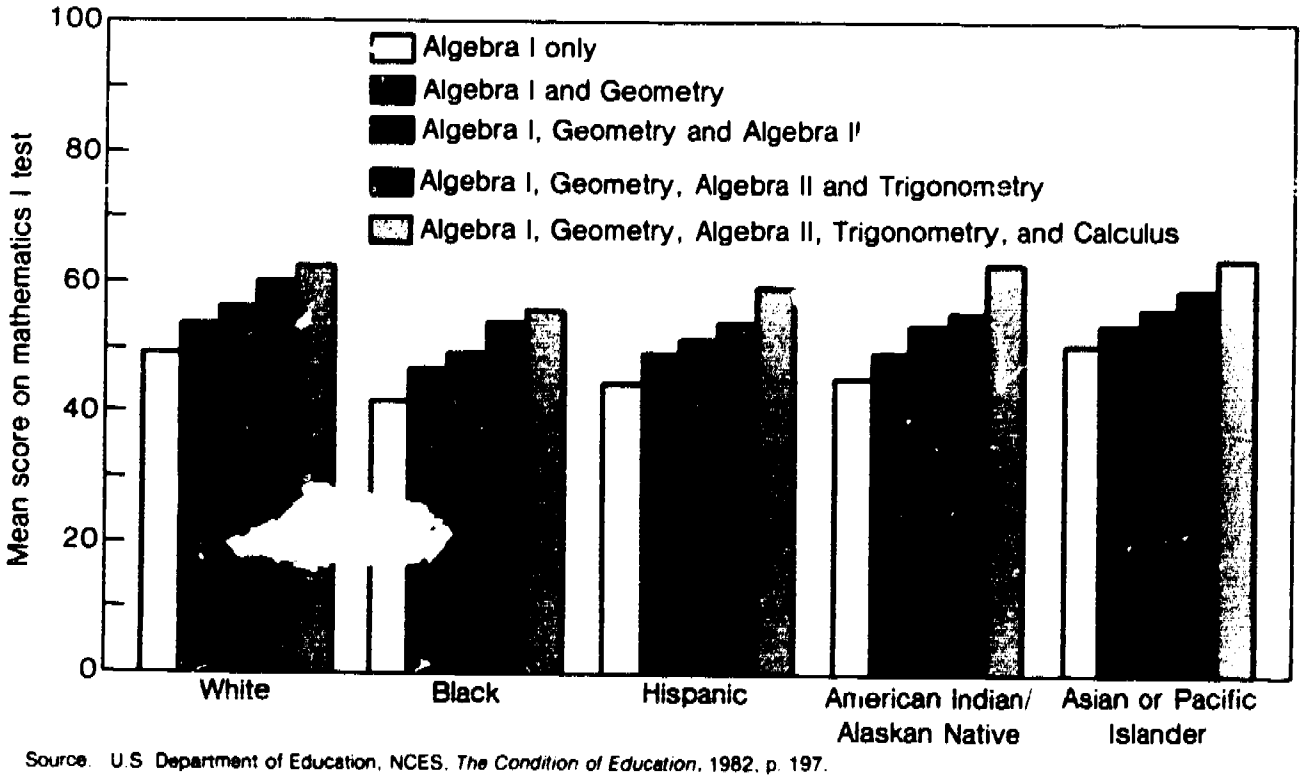
INDICATOR SYSTEMS

Whether indicators are single or composite statistics, rarely can a single indicator by itself provide adequate information about complex phenomena. More often, we try to construct a system of indicators to generate more, and more accurate, information about conditions.

An indicator system is more than just a collection of indicator statistics about a complex phenomenon, and it differs from a composite indicator that combines information to provide better understanding about one important aspect of that phenomenon. Ideally, a system of indicators measures *distinct* components of the system of interest, but also provides information about how the individual components work together to produce the overall effect. In other words, the whole of the information to be gained from a system of indicators is greater that the sum of its parts.

We are quite accustomed to using indicator systems in many areas of our lives. For example, when we want to know how comfortable we will feel on a summer day, we consider *together* separate indicators of the temperature, humidity, the air quality, and the probability of rain, rather than relying on any one of these indicators alone. No single indicator tells enough of the story by itself.

Although we currently lack complete educational indicator systems, we do have illustrations of how data from indicators might be used together to provide better information than we could get from any one indicator. For example, Fig. 5 presents together information about student ethnicity, coursetaking patterns, and achievement test scores. This juxtaposition of indicator data might suggest to policymakers that increased coursetaking could lead to increased achievement for all student groups.



Source. U.S. Department of Education, NCES. *The Condition of Education*, 1982, p. 197.

Fig. 5—Combined indicators of ethnicity, coursetaking, and achievement

In education, a complete indicator system should attempt to assess all of the relevant components of the educational system with a series of distinct indicators. For example, if we were only to have indicators of how well the system is meeting important goals such as achievement scores, we would lack other information—teacher quality, instructional processes, resources and materials—needed to judge its overall condition. Without a series of indicators that assesses all important facets of the schooling processes, we can neither understand the system's overall health nor determine the conditions under which a particular goal is met.

INDICATORS AND MODELS

The purposes of indicator systems are to measure the health and effectiveness of the education system and help policymakers make better decisions. These purposes imply that we select a set of indicators based on an understanding of which components of the educational system are critical to its health and which features signal important changes in its condition. We should also know how the various components of the system are related to one another. That way, particular combinations of statistics can be used to diagnose current and future conditions. For example, in automobiles, if the oil pressure drops *and* the engine temperature rises, we can reasonably assume that something is wrong with the system. Moreover, because we understand the relationship between these factors, that particular combination of conditions provides specific clues about what might be wrong.

Models Influence Selection of Indicators

Put another way, to properly specify which indicators should be a part of a system, we need a model of how the educational system works. If, for example, we began with the model shown in Fig. 6, we would construct an indicator system that included indicators of *each component* of the model. But a single indicator of each component of the educational system would still be inadequate. Within each component, we need indicators of all its most critical dimensions. For example, in the teacher quality component we would need more information than a single indicator could convey—e.g., a measure of teachers' formal qualifications. We would also want indicators of the essential features that define who teachers are, what they do, and the conditions in their schools that affect the quality of their teaching (Fig. 7). That way, the indicators in the system can provide essential information about each component of the model and illuminate how the various components relate to one another.

More than any other factor, the model chosen as the basis for selecting indicators will influence what information an indicator system will provide. But no single model of the educational process has gained universal or even widespread acceptance. This means that policymakers must attempt to assess the well-being and effectiveness of the educational system when we are not entirely clear (or in complete agreement) about what we want from it. We need to have some standard of educational goodness against which to judge *health* and *effectiveness*. But we have few clear-cut definitions of what each of these criteria should be. Unlike the economy where dollars stand out as a critical goal, we lack universally accepted standards against which to hold the educational system. The mutiple expectations we have for the schooling system and

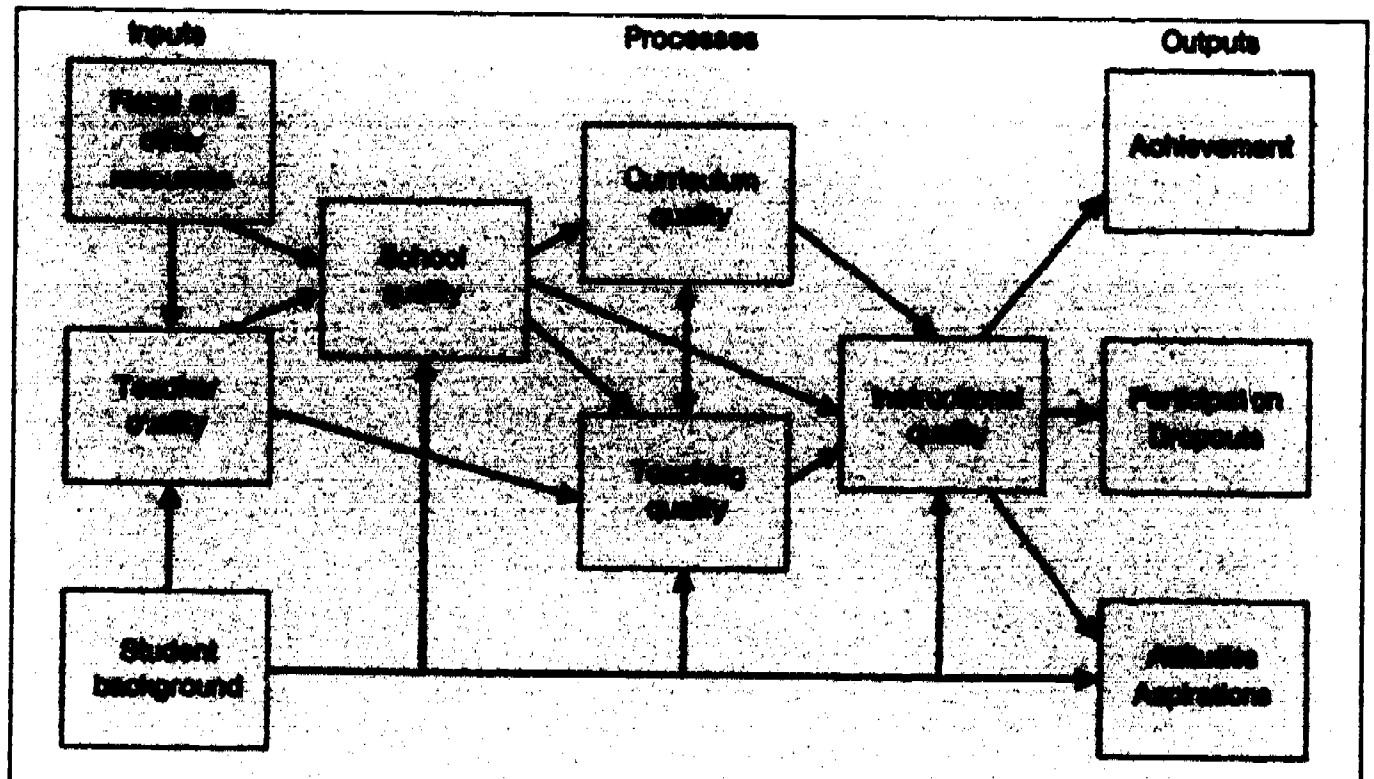


Fig. 6—A comprehensive model of the educational system

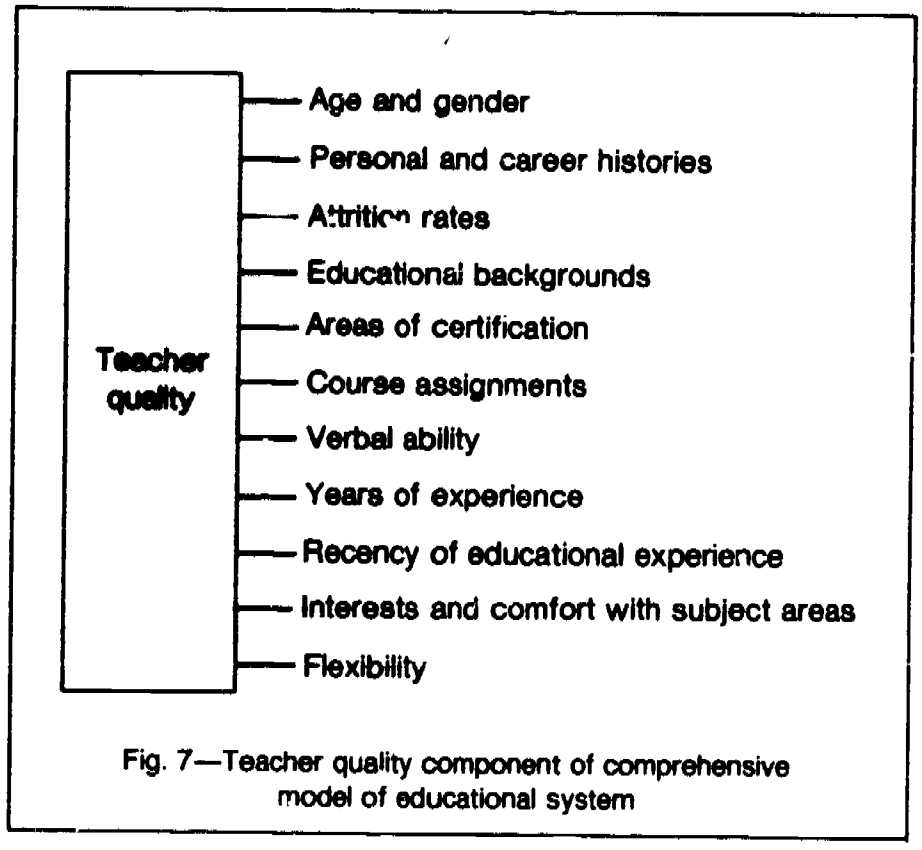


Fig. 7—Teacher quality component of comprehensive model of educational system

the limits of our knowledge about how it works make it difficult to know which statistics should be used to measure the educational system's health or effectiveness.

Functioning Under Uncertainty

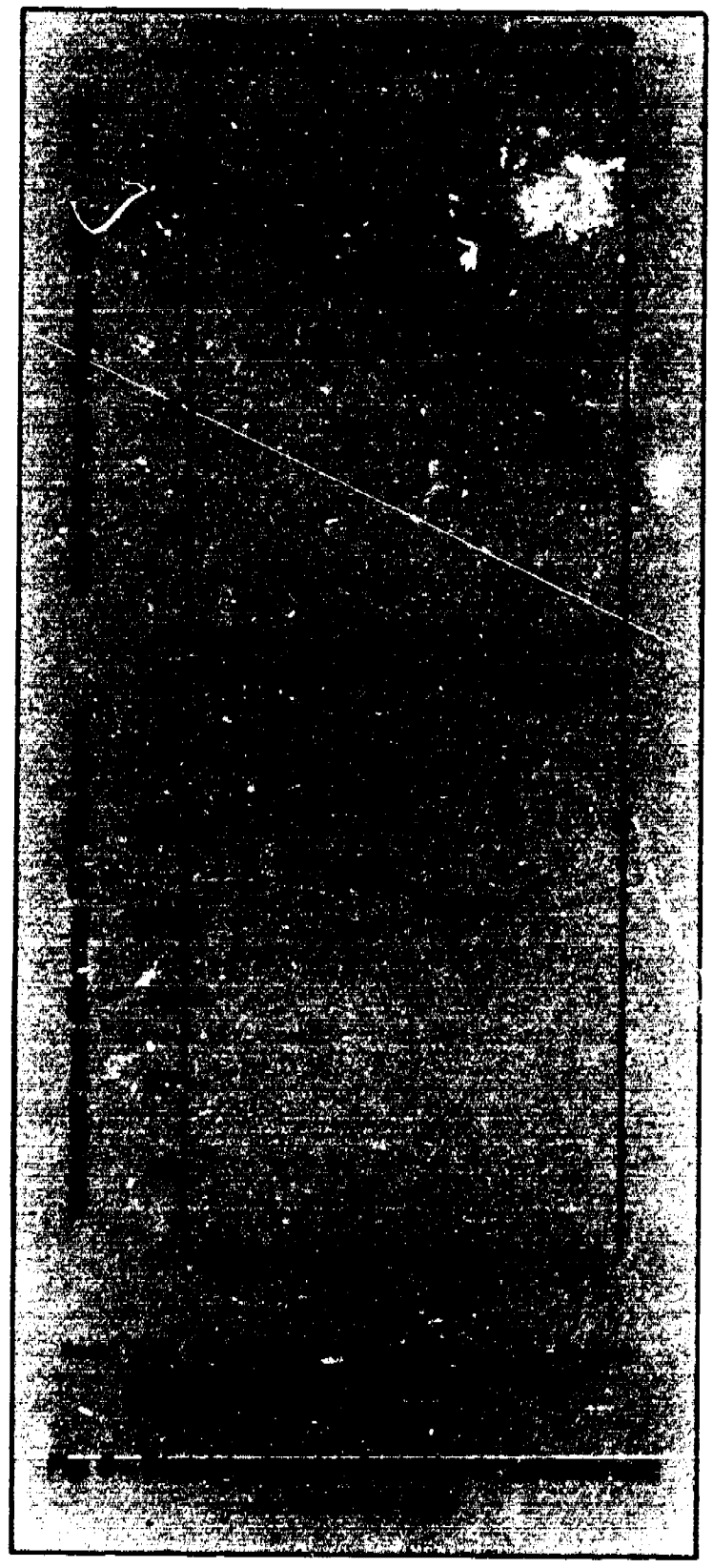
Although we currently lack a generally agreed upon model of the education system and, consequently, cannot construct a "perfect" indicator system to provide an overall assessment of the educational system's health or effectiveness, we do widely share views about some schooling outcomes. For example, most schooling constituents would argue that students' academic achievement and their willingness to stay in school until high school graduation are two highly-valued schooling objectives. If we have good measures of these two outcomes, we can use statistics from them as indicators of how well schools are accomplishing these desired results. Higher levels of achievement and lower numbers of dropouts could be used to signal a healthier, more effective system. Moreover, since we know that course enrollment and achievement are linked, we might use enrollment statistics as an indicator of the system's well being. The more students taking courses in subjects where we want increased achievement, the healthier and more effective we could declare the system.

We also have wide consensus about a number of schooling conditions that we want for their own sake. These, too, can be measured as indicators of well-being. For example, most of us want teachers to be well-trained, experienced, reasonably paid, and provided with manageable working conditions. Most would like children to be in small classes; to have ample opportunity to engage in problem-solving activities; and to have sufficient books, materials, and equipment at their disposal. Most agree that minorities and poor children should have opportunities equal to their more advantaged peers. For example, Fig. 8 illustrates the declining disparity in mathematics achievement between black and white students. The more of these conditions that exist, the healthier we might declare the system.

Recognizing the Limits of Indicators

Indicators can report what is, and we can use them to draw some conclusions about health and effectiveness. But indicators will only provide general information about schooling. We do not yet understand the educational system well enough to construct more precise measures of the system's condition or of the interaction among its components. No matter how broad the consensus about certain outcomes or conditions, without an agreed-upon model that integrates these features into the education system as a whole, statistics that measure them can provide limited information, at best.

These *caveats* point to the current limitations on educational indicator systems. However, they do not imply that we can not develop far better indicators than we currently have available. That we can surely do. But even better indicators must be interpreted with caution, and in the appropriate framework. Indicators can provide valuable information to guide the debate and dialogue about whether schooling is sick or well, effective or ineffective. But judgments about the health of the educational system can only be made by interpreting indicator data in the context of educational values and experience with schooling.



HOW INDICATORS ARE USED

In the last section, we described basic types of indicators and discussed the relationship between indicators and models of the education system. In this section, we discuss and evaluate some of the claims commonly made for indicator systems.

CLAIMS FOR AN EDUCATIONAL INDICATOR SYSTEM

The prospect of an educational indicator system has given rise to great expectations. Claims have been made that indicators will:

- report the status of American schooling; provide information needed to contrast schooling in different states and local districts, and to compare U.S. schooling with the education systems of other countries
- monitor changes over time
- explain the causes of various conditions and changes
- predict likely changes in the future
- profile the strengths and weaknesses of the system
- inform policymakers about the most effective ways to improve the system

Some claims appear quite achievable; however, others are clearly unrealistic. We'll consider each claim in turn.

Report and Compare the Status of Education

A complete set of indicators could summarize the human and material resources available to schools, the policies and practices that characterize them, and the results that they are achieving. For example, we might be interested in reporting and comparing the percentages of high school seniors who have completed various mathematics courses (Fig. 9).

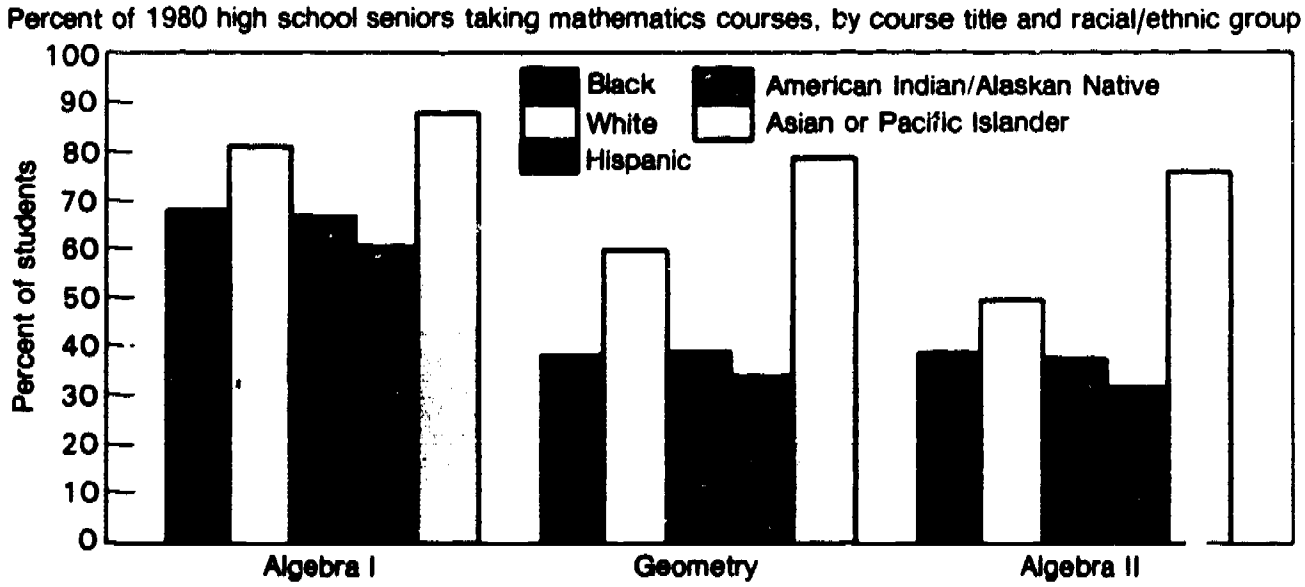


Fig. 9—Indicator reporting educational conditions

These indicators could be constructed to compare schooling in various local schools, states, regions of the country, and internationally. In fact, some existing, albeit limited, educational indicators (e.g., per-pupil expenditures, teachers' salaries, average class size, SAT and ACT scores) are currently used in just these ways. These are the indicators that have been incorporated by the Secretary of Education into his annual "Wall Chart" for comparing the education provided in various states (Fig. 10). Additionally, some states (e.g., California, Pennsylvania, Nevada) currently have systems that describe critical features of their own systems and compare statistics on a number of schooling features among local schools and districts.

STATE EDUCATION STATISTICS

State	Rank	Performance Outcomes: ACT (28 States) 1985 Score (1)	ACT 1985 Percent of High School Graduates Taking Test (Est.) (2)	ACT 1982 Score (3)	ACT 1982 Percent of High School Graduates Taking Test (Est.) (4)	ACT Score Change 1982-1985 (5)	SAT (22 States) 1985 Score (6)	SAT 1985 Percent of High School Graduates Taking Test (Est.) (7)	SAT 1982 Score (8)	SAT 1982 Percent of High School Graduates Taking Test (Est.) (9)	SAT Score Change 1982-1985 (10)
IOWA	Number	20.9	57.1	20.9	54.5	0.0	—	—	—	—	—
	(Rank)	(1) T	(10)	(2)	(15)	(15) T					
KANSAS	Number	19.1	62.1	18.9	60.8	+0.2	—	—	—	—	—
	(Rank)	(10)	(6)	(10) T	(8) T	(6) T					
KENTUCKY	Number	17.9	61.9	17.5	53.7	+0.4	—	—	—	—	—
	(Rank)	(19)	(17) T	(23) T	(16)	(2) T					
LOUISIANA	Number	16.5	60.7	16.7	60.8	–0.2	—	—	—	—	—
	(Rank)	(27)	(8)	(27)	(8) T	(25) T					
MAINE	Number	—	—	—	—	—	888	51.5	890	46.4	+ 8
	(Rank)						(11)	(12)	(8)	(18)	(17) T
MARYLAND	Number	—	—	—	—	—	910	51.3	889	50.3	+21
	(Rank)						(6)	(13)	(9) T	(11)	(3) T
MASSACHUSETTS	Number	—	—	—	—	—	906	66.8	896	65.6	+10
	(Rank)						(8)	(2)	(11) T	(2)	(10) T
MICHIGAN	Number	18.9	51.9	18.7	51.4	+0.2	—	—	—	—	—
	(Rank)	(11) T	(17) T	(12) T	(18) T	(6) T					
MINNESOTA	Number	20.2	29.5	20.2	26.9	0.0	—	—	—	—	—
	(Rank)	(3)	(28)	(3)	(28)	(15) T					
MISSISSIPPI	Number	15.5	64.2	15.5	74.4	0.0	—	—	—	—	—
	(Rank)	(28)	(2)	(28)	(1) T	(15) T					
MISSOURI	Number	18.8	48.9	18.7	45.3	+0.1	—	—	—	—	—
	(Rank)	(14) T	(21)	(12) T	(23)	(12) T					
MONTANA	Number	19.5	52.1	19.5	49.5	0.0	—	—	—	—	—
	(Rank)	(6)	(16)	(6)	(20)	(15) T					
NEBRASKA	Number	19.7	62.9	19.9	73.0	–0.2	—	—	—	—	—
	(Rank)	(4) T	(4)	(4)	(2)	(25) T					
NEVADA	Number	18.5	43.6	18.3	44.5	+0.2	—	—	—	—	—
	(Rank)	(17)	(24)	(18)	(24)	(6) T					
NEW HAMPSHIRE	Number	—	—	—	—	—	939	57.0	925	56.4	+14
	(Rank)						(1)	(6)	(1)	(6)	(14) T

Source: U.S. Department of Education, 1986.

Fig. 10—"Wall Chart" indicators comparing states

Comparisons must be based on universal features

Although *reports* that describe school conditions might include some features of schooling that vary from place to place, *comparisons* must be based on components of schooling that exist everywhere. For example, it would be of little use to compare school systems on the quality of programs in agriculture because programs in agriculture are not common to all schools. In fact, one of the flaws in the current "Wall Chart" is that its indicator of student achievement is based on scores on college entrance exams that are taken by only a fraction of the nation's students—and this fraction varies from state to state.

Comparisons require consistent definition of measures

Even more important, making appropriate comparison among states and local schools will require that indicators be defined and measured in the same ways. Currently this is not the case, even for the most basic features of schools. For example, methods for computing school attendance rates vary considerably from state to state: Some states count only those students actually present at school as attending; in other states, students staying home with an "excused" absence are counted as attending. Student achievement provides another example of how inappropriate comparisons can distort information. Although most states have systems in place to assess students' achievement outcomes, the specific measures they use vary considerably. If we were to use statistics produced by different definitions or measures to compare states or schools, we would draw very unreliable and, perhaps, quite erroneous conclusions.

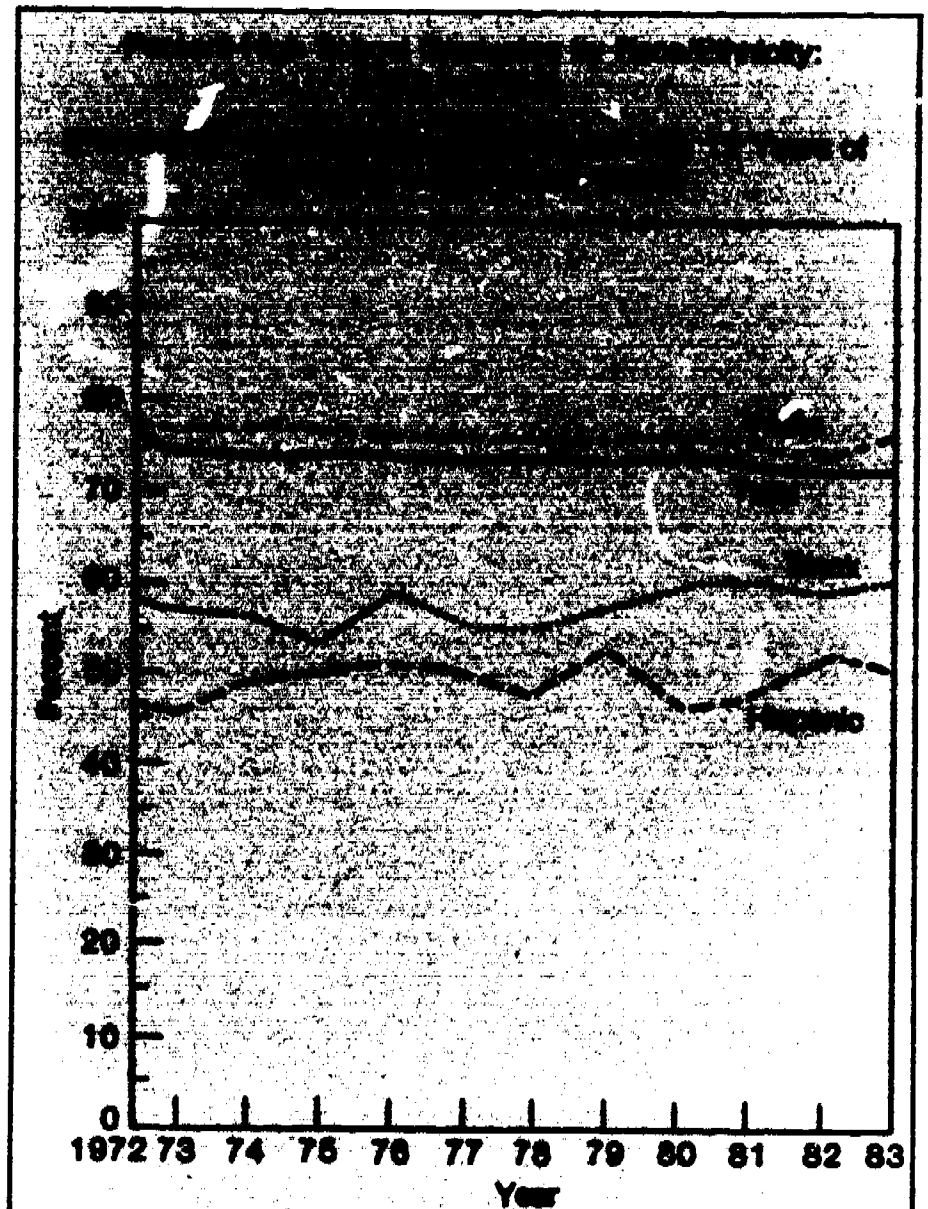
Track Changes Over Time

An indicator system that collects, analyzes, and reports data about schooling *at regular intervals* will show changes and trends in the educational system over time. Currently, we use many educational statistics in this way. We plot high school graduation rates and the ups and downs in SAT scores each year; we track changes in achievement from one administration of the National Assessment of Educational Progress to another; we track the variations in the number of college graduates choosing teaching as a career, to name just a few (Fig. 11).

Indicators also allow us to track changes that follow significant policy decisions or that parallel other changes in the education system. For example, states that experience major shifts in sources of funding could track changes in schooling that follow such shifts (Fig. 12). States or districts might be interested in tracking dropout statistics following increased high school graduation requirements or the imposition of competency examinations.

Tracking establishes trends, not causes

Tracking these changes does not, of course, imply that we have established cause and effect since in natural settings like schools, many factors may cause the trends we observe. Tracking changes may stimulate hypotheses about causes and identify areas for further investigation, but the existence of a trend does not *in itself* permit us to draw inferences about causality.



Source: Bureau of the Census, *Current Population Reports*.

Fig. 11—Indicators used to track changes over time

Tracking requires enduring features, constant measures

Obviously, if indicators are to track changes, they must measure enduring features of the system, and measures themselves must remain constant over time. Ensuring such constancy means paying particular attention to how the measures are defined. For example, we might want to compare teacher-student contacts over time, i.e., how many children teachers are responsible for and how many adults children have available to them in school. One way of defining this feature of schooling is *the proportion of full-time teaching positions at a school to its student population.* Unlike class-size or te-ching load, this definition could lead to reliable statistics even if scho organization changed. For example, if a school switched from self-contained classrooms to a departmentalized system, or to large team-taught classes, this definition would still capture the important ratio of teaching adults to students in the system.

State	Current expenditures per pupil, 1970-71 [illegible][1]	Current expenditures per pupil, 1980-81 [illegible][1]	Current expenditures per pupil, 1982-83[2]	Percent change (in 1982-83 dollars) 1970-71 to 1980-81	Percent change (in 1982-83 dollars) 1980-81 to 1982-83
United States	[illegible]	[illegible]	[illegible]	25.2	4.6
Alabama	[illegible]	[illegible]	2,177	50.5	−3.2
Alaska	[illegible]	[illegible]	[illegible]	73.1	18.8
Arizona	[illegible]	[illegible]	[illegible]	37.2	−1.4
Arkansas	[illegible]	[illegible]	1,971	[illegible]	2.4
California	[illegible]	[illegible]	2,738	25.1	−2.5
Colorado	2,079	[illegible]	3,171	[illegible]	3.9
Connecticut	2,910	3,176	[illegible]	21.7	14.6
Delaware	[illegible]	3,412	[illegible]	[illegible]	1.3
District of Columbia	2,912	[illegible]	[illegible]	[illegible]	9.2
Florida	[illegible]	[illegible]	[illegible]	20.0	9.0

[1]Adjusted for inflation using the Consumer Price Index computed on a school-year basis.
[2]Preliminary data.

NOTE: Details may not add to totals because of rounding.

SOURCE: U.S. Department of Education, NCES, [illegible]

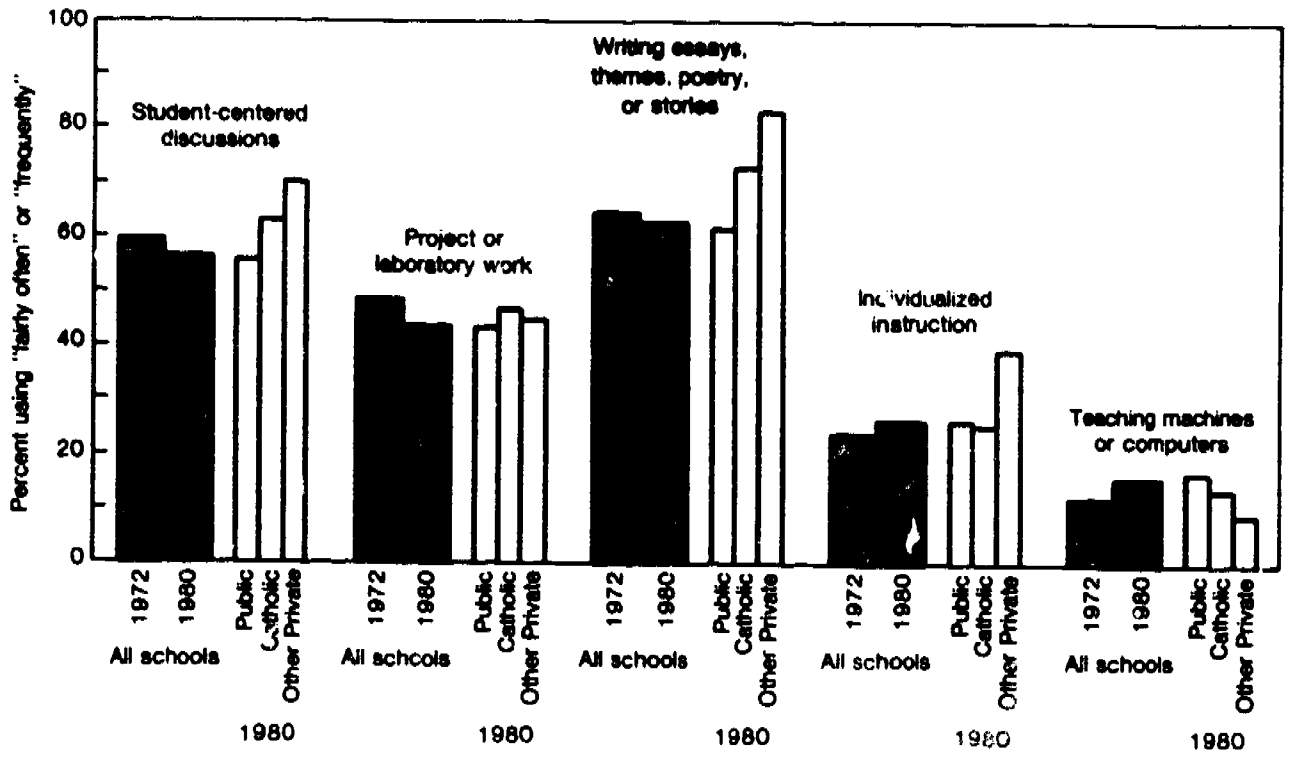
Fig. 12—Indicators used to track changes during period of changing school revenue sources

Similarly, if defined in generic ways and measured consistently, classroom instruction indicators could document variations in instruction both over time and among school types (Fig. 13).

The definitions of indicators must remain constant over time even if a better means of measuring the same phenomenon is perfected, or comparisons over time will be rendered invalid. This requirement will probably be a dilemma for educational indicators since, as we will discuss in greater detail below, few "good" measures of even the most important features of schooling are currently available. However, there are solutions to this dilemma. For example, as new measures are developed they can be used concurrently with older ones to provide statistical bridges or "crosswalks" and preserve the continuity of trend data. In that way, better indicators can evolve over time.

Predict Future Performance

One of the most appealing claims made for indicators is that they will serve as an early warning system, alerting policymakers to problems in time to intervene. But, predicting future conditions or events about complex social systems like education from the movement of indicators is difficult because we know little about what actually causes changes in these systems to occur. Economic indicators provide a parallel example of this difficulty: we have not been terribly successful at predicting the future because we do not fully understand the causes of changes in the economic system. In education we may fare even less well.



Source: U.S. Department of Education, NCES, *The Condition of Education*, 1982, p. 83.

Fig. 13—Indicator of enduring school features that can be compared over time and among school types

Indicators as harbingers of change.

The most straightforward way that indicators can predict future changes is when they collect data directly relevant to the future. For example, indicators that measure birth rates or population shifts can provide quite accurate projections of school populations in the future. Similarly, indicators that assess career intentions of high school seniors can provide an early warning of possible teacher shortages or surpluses in the future.

Leading indicators can predict accurately

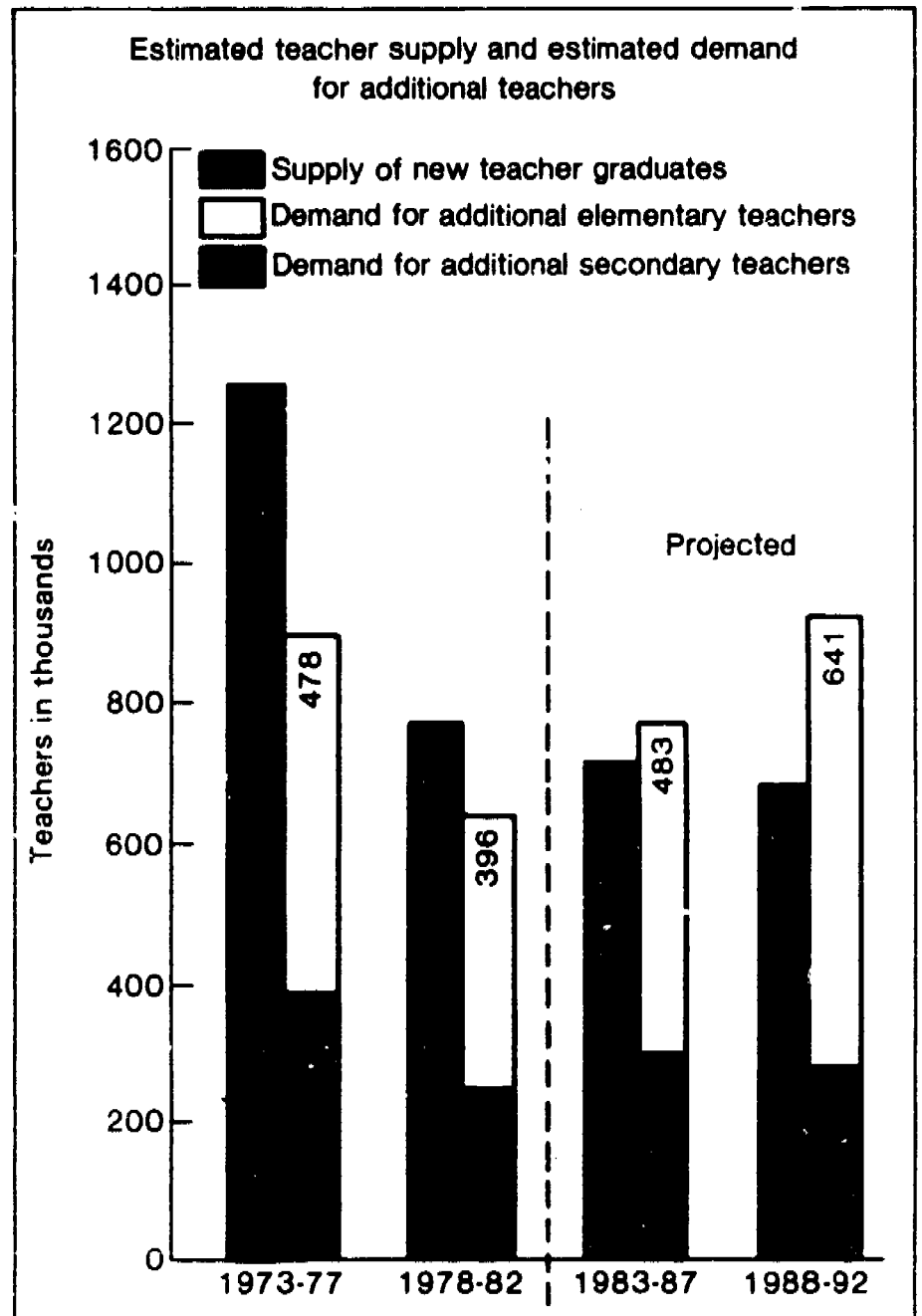
To the degree that theory and research have established links between certain indicators and other events in the educational system, indicators can probably be used to predict quite accurately. And, we do know some things. For example, we know that the more classes secondary students take in subjects like science and mathematics, the higher they are likely to score on achievement tests in those subjects. So, if scores on an indicator of science and mathematics courses participation went up, we could predict an increase in mathematics achievement scores with considerable confidence.

This type of prediction requires that some indicators qualify as "leading" indicators, i.e., when they change, other known changes will follow. In the economic sphere, the Index of Leading Economic Indicators has such predictive value. With this economic index we can predict, for example, that a decline in raw materials portends a subsequent decline in

manufacturing output; or, a reduction in inventories will lead to a reduction in consumer sales.

Patterns of change support less certain predictions

As with most economic indicators, howeve , a less certain type of prediction is all we should expect from an educational indicator system—predictions based on observations of trends in the past, but with no absolute certainty about causes (Fig. 14). Over time, patterns of changes among educational indicators may be noted—for example, when teachers' salaries increase, we may note a subsequent increase in the number



Source: U.S. Department of Education, NCES, *The Condition of Education*, 1984.

Fig. 14—Indicator used to project trends

of college graduates entering teaching. When such trends appear, they can be used to project conditions in the future. Although we cannot relate conditions in a causal way (e.g., more students are entering teaching *because* salaries increase), we could predict the former on the basis of the latter.

Predictions based on indicators are likely to become increasingly precise as patterns recur over time and if combinations of leading indicators point to the same future events. But, as with economic indicators, we will undoubtedly discover that many anticipated changes never come to pass and that unanticipated ones spring up repeatedly. Indicators may improve our ability to see what's on the horizon, but we should not expect them to provide a perfect description of the future.

Explain Conditions and Changes

Many hope that, in addition to tracking and projecting changes, indicators will be able to provide explanations for why things happen as they do—to determine what factors caused particular changes and link particular policies with their effects. The data from indicators might indeed document such links, but only when the connections are obvious or if prior theory and research have established the relationship between one component of the system and another. For example, we might note a sharp upward trend in an indicator of mathematics teachers' qualifications following the institution of a loan-forgiveness program for students entering these fields. Of course, it would be tempting in this case to conclude that the policy *caused* the change in the number of qualified teachers. Such an inference might be warranted, but, if so, it would be because of the logical connections we make between the two events and other information we may have at hand. The fact that the indicator changed would not, in itself, prove the connection between the two events.

Without prior research, indicators can't establish cause and effect

This rather subtle distinction is critical because, invariably, policymakers and researchers will want to interpret the findings of a monitoring system in an *explanatory* fashion. That is, if one indicator goes up or down, other changes in indicators will be scrutinized for an explanation, even when prior research has *not* established links between indicators. For example, if there were to be an increase or decrease in an indicator of student achievement at the same time that there were changes in indicators of educational spending, teacher qualifications, or school organization, it would be tempting to use these other changes to "explain" why students' scores moved in the direction they did.

Of critical concern to policymakers is determining whether their decisions have the intended effect. An indicator system can track effects when relationships are already known from prior research. And, even when effects are not known, indicator data can provide clues about how components of the system are related, and they can suggest possible relationships for researchers to explore. Unfortunately, few conditions in education have causes that are clearly understood. Consequently, we risk using the information from indicators inappropriately and misspecifying causes when there are none or inferring the wrong direction of influence.

An example may help to clarify this point. A recent report[1] documented that at least 79 theories have been advanced to explain the steady decline of SAT scores between the early 1960s and 1980 (including drugs, parental neglect, poor teacher training, and the birth-rank of test-takers in various years). For an indicator system to *establish* the cause of this decline, we would need to include in the system indicators that measure all the factors suggested by the various theories so that we could rule out all the irrelevant factors at the same time that we pointed to those actually contributing to the decline.

If including all of the factors that might explain achievement declines seems unwieldy, we must remember that we would also need to include *all the possible causes of all the other conditions or events we hope to explain*—student dropouts, teacher shortages, special problems of girls and minorities—and some controls to establish which particular change was responsible for a given effect.

Indicators can provide clues

However tempting the conclusions, an indicator system can not provide all the necessary conditions and controls to determine the causal relationship among these events. Indicators may call to our attention the fact that two events occurred at the same time—that is, that they are correlated. However, they provide no grounds for concluding that one event caused the other. An indicator system can track causes and effects when relationships are already known from prior research. And indicators can provide clues about about how components of the system are related; thus they can suggest possible relationships for researchers to test using more controlled procedures. But we cannot expect an indicator system itself to explain why things happen as they do.

Profile Strengths and Weaknesses of the System

We have considered expectations that indicator systems will report conditions, track changes, predict the future, and explain why particular changes occur. Additionally, an educational indicator system is expected to profile the strengths and weaknesses of the educational system and to evaluate how well it functions. In this way, indicators are anticipated to provide a public "report card," and hold the system accountable by permitting public scrutiny of the performance of the system. This expectation, too, is likely to be fulfilled in a far more limited way than many might hope.

Need reference points to judge system

Profiling the strengths and weaknesses of the educational system poses the fundamental problem for policymakers raised earlier: the difficulty of judging the *overall* health of a system when we are not entirely clear—or in complete agreement—about what we we want from it. The problem is very basic: For a statistic to provide information about the strengths or health of a system, we need to have a reference point for it—in this case, we need to have some standard of educational "goodness." At the very least, we need a sense of whether more or less of particular schooling characteristics signals that conditions are "better" or "worse," "healthy" or "sick." If we know which direction is healthy, we

[1]Wharton, Y. L. *List of Hypothesis Advanced To Explain the SAT Decline*, 1977, College Entrance Examination Board, N.Y.

can assess the relative conditions of the system over time and from place to place.

However, as a society we have not precisely defined schooling goals. For example, we know that we want schools to develop students academically, socially, personally, and vocationally. But we have no clear-cut definitions of what each of these means, or how we rank these goals. We also want schools to keep children safe and comfortable, even happy. But it's not clear how much of children's happiness we would be willing to sacrifice for their academic progress, or vice versa. Without such standards, it is exceedingly difficult to define strengths or weaknesses or to choose statistics to measure either one.

Because we lack basic reference points, we cannot use an indicator system to make an overall judgment about the strength or weakness of the educational system. The best we can do is to measure those outcomes—such as achievement and graduation rates—and those schooling conditions—such as well-trained teachers and manageable working conditions that we all agree are desirable in themselves. We might reasonably conclude that the more of these conditions that exist, the stronger and healthier the system is.

In the end, however, educational indicators will probably function more like the weather report than like sports statistics or even economic indicators. We probably will find few clear-cut good or bad conditions, few unequivocal ups and downs.

Inform Policy and Practice

Finally, an indicator system is expected to inform policymakers and educators about policy and practice—to point to the types of changes that will be appropriate and effective interventions. Moreover, as educational reforms are implemented, the indicator system will be expected to assess their contribution to better schooling. This is a tall order for indicators, particularly for an indicator system intended to serve national as well as state and local interests. If state and local policymakers expect to use indicators to track the implementation and consequences of policy, they will probably need to develop systems of state and local indicators designed specifically to reflect their particular conditions and concerns.

Limits of national indicators

A national indicator system is likely to be guided by federal articulations of educational goals and consist of measures that will be appropriate for national descriptions and comparisons (i.e., measures of characteristics commonly found across states and schools). But most decisions about the conduct of education are made in states, local districts, and individual schools. State and local policymakers may or may not incorporate national recommendations into policy goals and actions for their schools. If they do, the specific forms they take are likely to vary considerably from place to place. Currently, many states have adopted curriculum policies, policies regarding the use of instructional time, and policies regarding teacher qualifications and compensation. But these policies exhibit considerable variation in substance, extensiveness, and their implementation time-frames across states and localities. Indicators

that meet the criteria for a national system are unlikely to focus on the multitude of targets at which these policies are aimed.

Further complicating the problem of developing national indicators that could be useful for informing state and local policy is the rapid pace of change within the educational policy environment. Simply keeping track of what state and local policies have been adopted, their intended effects, and the extent of their implementation would be a major data collection and analysis effort. Constructing and implementing indicator of all of them as national indicators would be impossible.

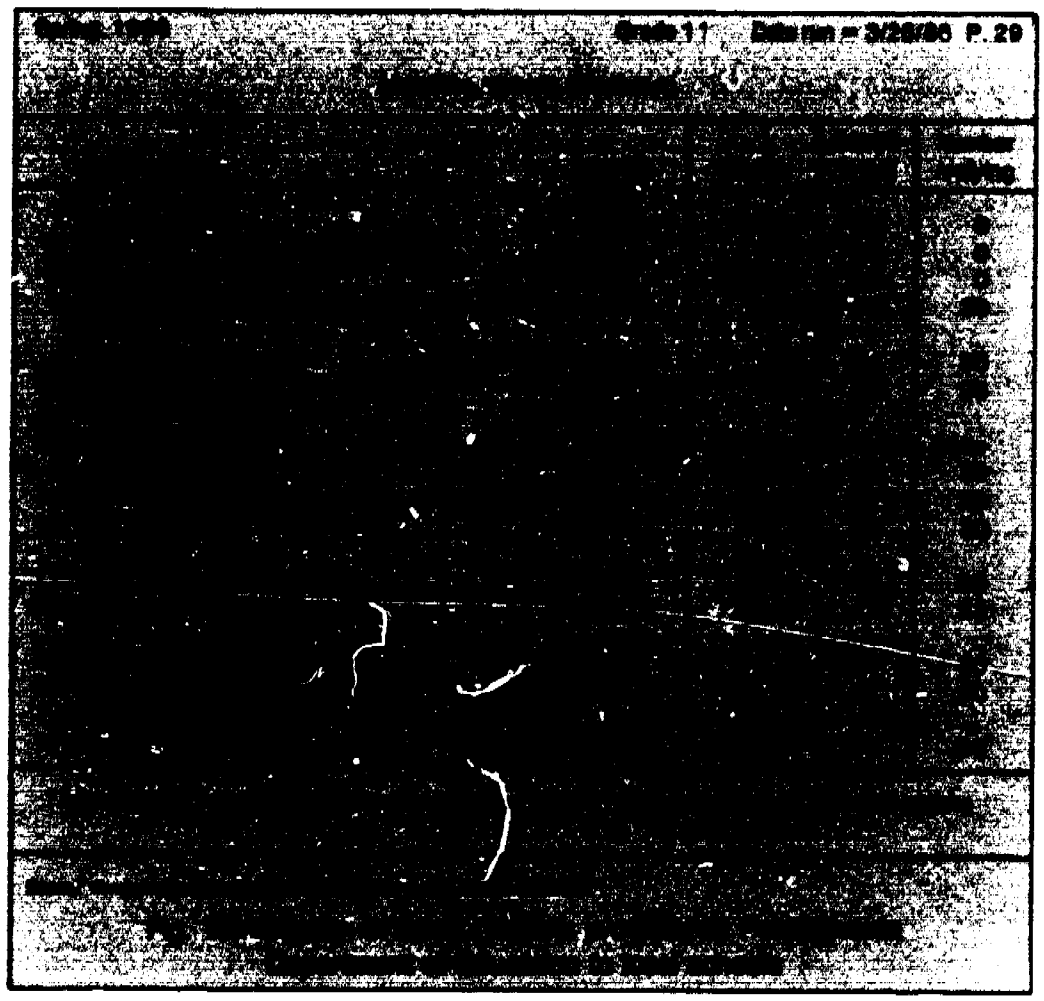
As a result, when it comes to providing data to inform policy decisions, the most we can probably expect from national indicators is that they provide information about what generic types of policies (e.g., mandates, capacity-building, incentives) or general categories of policies (e.g., curriculum policies, teacher policies) are being implemented in various states and local districts and what general types of subsequent changes seem to be associated with them.

Potential of state and local indicators

Indicators developed by states and local districts could, on the other hand, be quite useful for monitoring particular state and local policies and practices. At these levels, indicators could be designed to reflect specific conditions and concerns that have been the target of policy initiatives. For example, scores on the student achievement measures used by a state could constitute a state-level indicator of important outcomes that would provide information more relevant and specific to a state's particular achievement goals. This more specific information could supplement the more general achievement information supplied by a national indicator. The achievement and participation of sectors of the student population of particular interest could be tracked with state- and locally-developed indicators as well. For example, an indicator that specifically charts the progress of non-native English-speaking children of migrant workers might be of interest in states like Florida and California. Knowledge and competencies that have been designated as targets for improving teacher quality could be incorporated into state-specific teacher-qualifications indicators. Moreover, if states collect these data, they can then provide information back to local schools and districts (Fig. 15).

Indicators that focus on specific conditions, educational practices, and outcomes of interest could indeed provide policymakers with clues about what changes have occurred following the implementation of policies aimed at those factors. Caution must be exercised here, of course. Indicators that are too specific to current concerns may be of little use over time. As with national indicators, indicator developers at the state and local levels should take care to develop measures of central features of the system or those targets of policy initiatives that are likely to be of interest even if immediate policy priorities shift.

Of course, expecting state and local indicators to signal what specific policies and practices will work toward specific goals raises all of the difficulties of indicators noted earlier. Indicators will report conditions; and, as new indicators are developed, they certainly will result in far better information than educational statistics currently provide.



Indicators will generate knowledge about the educational system that will fuel the ongoing debate about sc' ooling. They will influence how policymakers and the public think about schooling and provide useful clues about promising policies. But even as we develop better indicators, they will neither identify specific causes, nor provide unequivocal solutions to problems. In the end, interpretations of what indicators mean and decisions about what policies should be implemented will be influenced by values as well as by knowledge. Indicators can not remove fundamental decisions about schooling from the political process—from the reflection and debate among policymakers and the public ultimately responsible for its healthy functioning.

Indicators in that broader policy context is the subject of the next section.

INDICATORS IN A POLICY CONTEXT

FIVE POLICY ISSUES

We have seen that decisions about desired outcomes and conditions will determine the nature of any indicator system. In large part, these decisions will be political—that is, influenced by the concerns of specific groups, by public opinion, and by educational goals at different levels of the system. In this section, we consider the broader political context in which any indicator system must be implemented and speculate about how this context will shape an indicator system. We will focus on five basic issues: (1) the level of information to be collected, (2) the challenge of making comparisons, (3) the costs and benefits of extensive indicator systems, (4) the political pressures that the existence of an indicator system will bring, and (4) the vital question of who makes the design decisions for any indicator system.

What Level of Information Should the System Collect?

One source of tension about what and how to monitor stems from the divergent information needs of local policymakers and practitioners, state policymakers, and federal education officials. As noted earlier, federal officials will undoubtedly be most interested in looking at characteristics that are evidenced across states and local districts since these will permit both national description and state-by-state comparisons. State-level policymakers will probably want indicators capable of tracking the effects of state policies. Local decisionmakers would probably like site-specific data that would be useful for improving particular programs. Equally vexing differences among levels will surely plague decisions about methods of indicator data collection, analysis, and reporting.

Top-down or bottom-up?

A fundamental decision is to what extent a national indicator system should be *top-down* (shown in Fig. 16) or *bottom-up* (shown in Fig. 17). A top-down system would consist solely of indicators defined at the national level and measured commonly across schools and districts. A bottom-up system would begin with indicators deemed relevant to states and localities. To the extent that various states had similar information needs and could agree upon indicators, these would constitute a national system. A combination of top-down, bottom-up would proceed from the national and state/local levels simultaneously, and would require a great deal of cooperation and coordination among them.

We can illustrate the difference between these two kinds of systems by comparing economic indicators and weather indicators.

Economic indicators are a top-down, national system. They are defined and collected nationally. Few states or local communities have well-developed local economic indicators, but, to the extent they are useful, national indicators broken down by states and regions inform local policymakers and businesses. This system is appropriate since economic policy is largely a national concern. For example, few local banks set

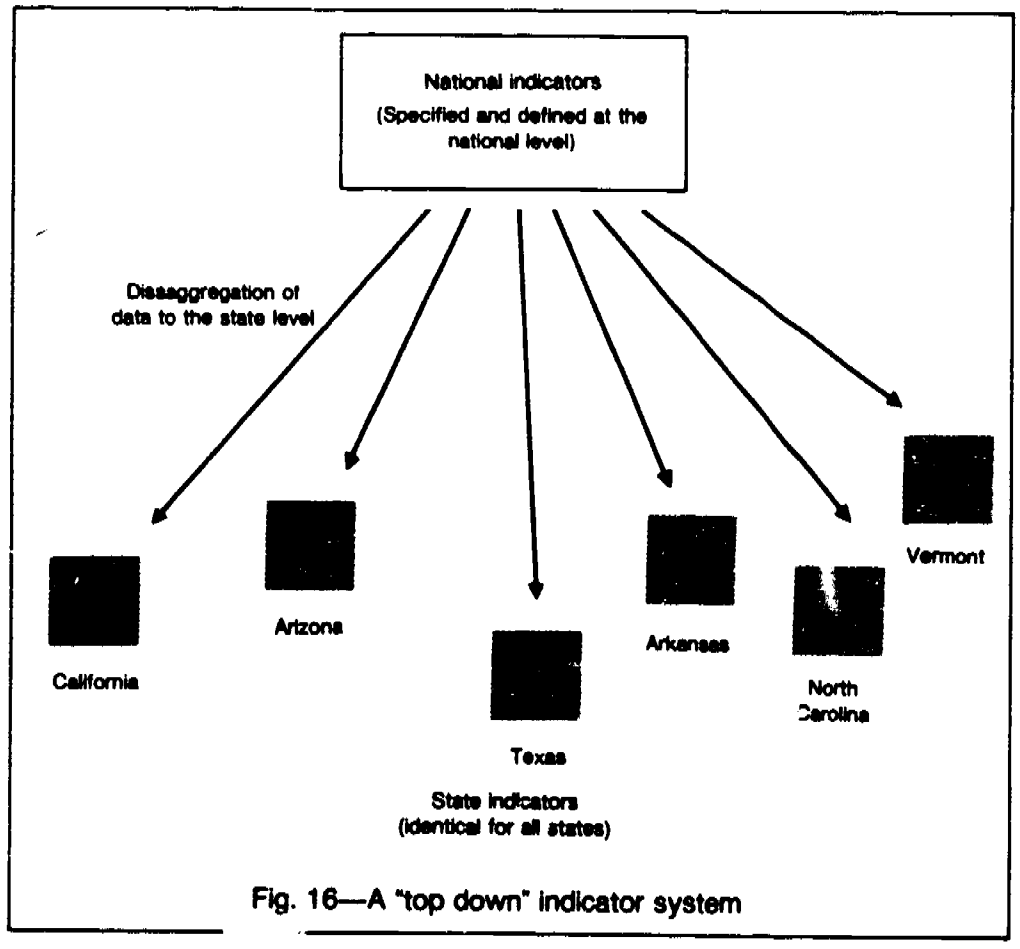


Fig. 16—A "top down" indicator system

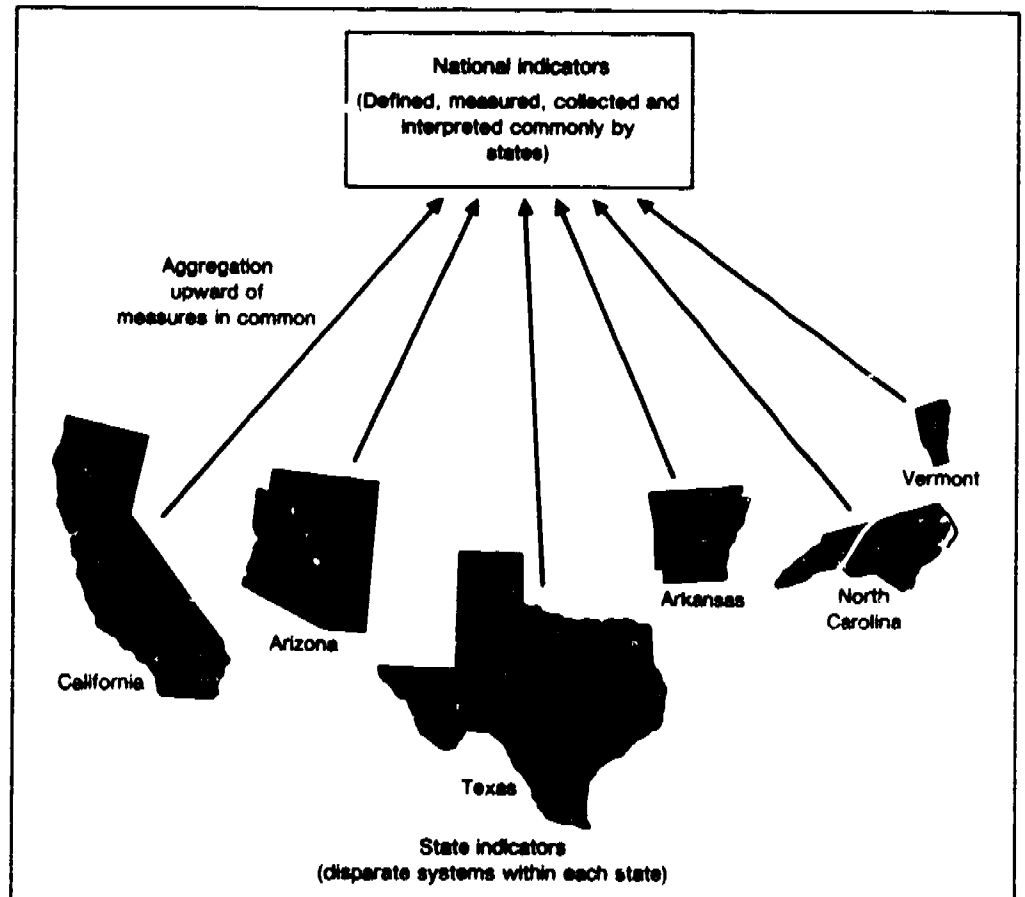


Fig. 17—A "bottom up" indicator system

their own interest rates independent of moves by such federal bodies as the federal reserve or major banks with national constituencies.

The weather report, on the other hand, appears to be a largely bottom-up system. Local reporting agencies use agreed-upon definitions of indicators and common measures, and report local data to central agencies. Local weather indicator data are reported at the regional and national levels, but the data are rarely aggregated. That is, the national weather report does not include the nation's average temperature, humidity, or rainfall. Some conditions specific to a local region—smog or storm warning, for example—may never or only in extreme circumstances, be a part of national or regional weather reporting. On the other hand, local data provide specific information that is needed for local decisionmaking—declaring smog alerts that restrict or curtail industrial activity and driving in major cities, or issuing hurricane advisories along the gulf coast.

The weather system is not entirely local; weather satellites are perhaps the most centralized data-gathering operations imaginable. However, the largely bottom-up approach to weather reporting is clearly appropriate given the great variety of weather conditions across the country and the relative uselessness of national averages to any one locality.

Which approach is more appropriate for schooling? Both sides of the argument have their energetic supporters. States and localities want to preserve their identities and gather information relevant to their particular concerns. While the federal government is interested in helping states acquire these data, its primary concern is for the condition of the country's educational system as a whole. Consequently a major emphasis is on data that will allow generalizations across the nation and comparisons among states. Realistically, for any indicator system to work, all those who will participate in the system—either by supplying or using data, or by paying for the system—must perceive it as an information source that serves their needs and interests. Achieving this harmony of motivation is likely to require cooperation and compromise on all sides.

How Can States and Schools Be Compared Fairly?

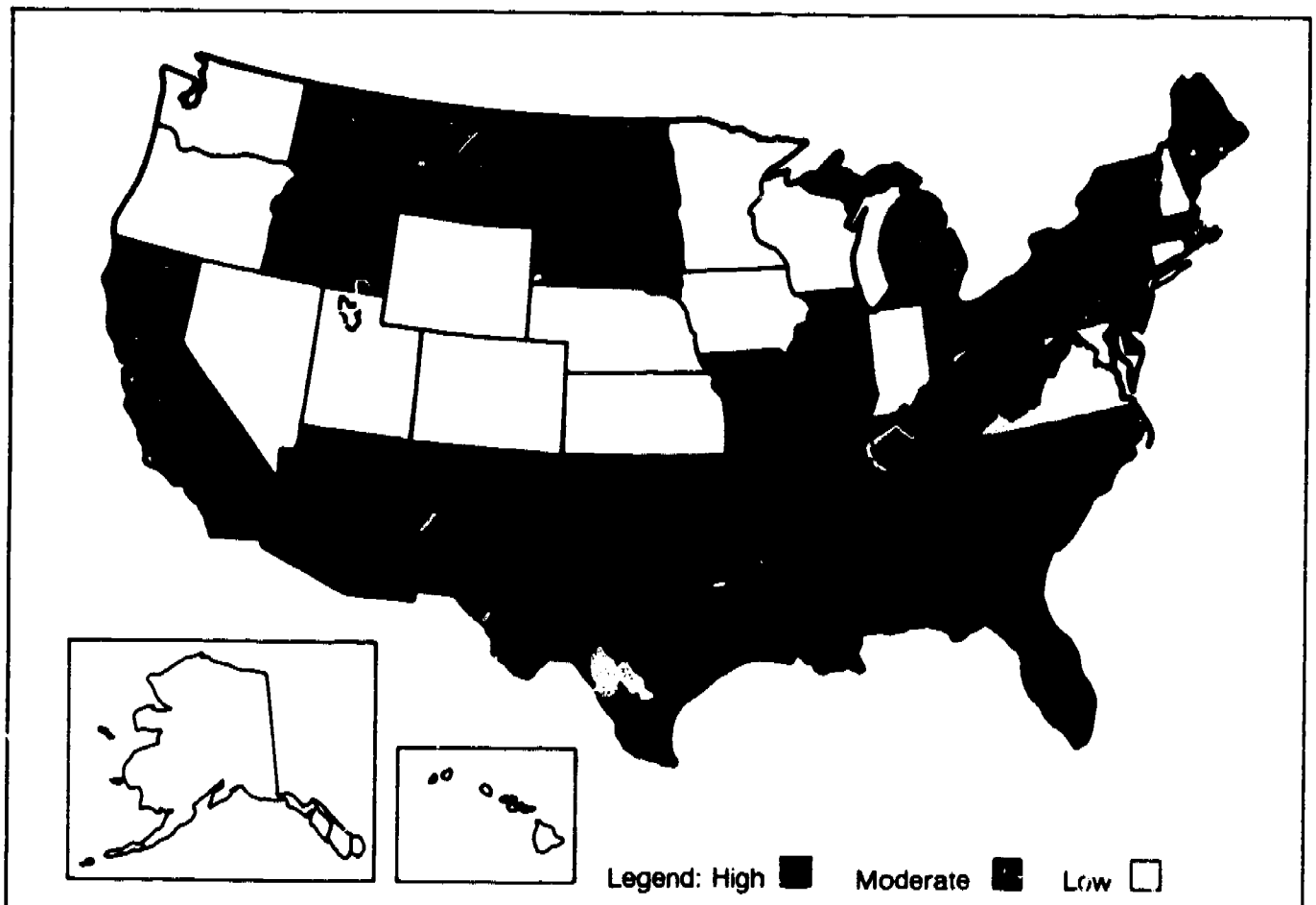
We have already noted that one function of an indicator system is to compare states, schools, teachers, and students. But the ability to make such comparisons raises a difficult political issue: How do we build into a system of indicators the reality of differences and account for them when we make comparisons without institutionalizing lower expectations for some schools and their students?

Can indicators reflect differences while promoting equity?

Not all states, schools, teachers, and students start out even. Some systems face greater difficulties than others in educating; some students face greater difficulties becoming educated. We know that achieving desired educational outcomes will be more difficult for schools with large numbers of poor children, minority children, or non-English-speaking children. We know that states and regions with high levels of transiency will face more problems in reaching educational goals. States with particularly depressed economies or high levels of unemployment will find it more difficult to generate adequate resources for schooling.

Figure 18 presents some of these differences among states. We want all children, regardless of race, gender, religion, physical or mental handicap, or social standing, to have a fair opportunity to succeed. So we expect schools to *overcome* the difficulties they face. But the differences among them mean that, although two states or schools may achieve the very same results, their actual *accomplishments* might vary considerably.

Out of a sense of fairness to states and to local schooling systems, indicators should be sensitive to these variations. Yet, out of a sense of fairness to the children of various circumstances, indicators should account for the distribution of educational outcomes and processes among various student groups—indicators that will monitor the system's equity. Indicators that include "adjustments" for the educative difficulties various schools face—for example, different expectancy ranges in



Source: U.S. Department of Education, NCES, *Indicators of Education Statistics and Trends*, 1985.

Fig. 18—Index of requirements for special educational services

achievement scores for schools serving students of different socioeconomic backgrounds—run the risk of institutionalizing low expectations for some students. On the other hand, indicators that don't account for these important differences would place an undue burden on many schools when comparisons are made. They could obscure real achievements of schools with the most difficult circumstances and inflate the accomplishments of schools working under far easier conditions.

There are possible solutions to the comparison problem. One approach would be to simultaneously emphasize educative difficulties and the goal of equity. Indicator data might document both the *overall difficulties* a state or district might face and the *distribution of school conditions and attainments* among student groups within these states and districts. If disparities among groups were also a focus of attention, the difficulties schools and districts face might be less likely to lead to lower expectations for the least advantaged students. Constructive approaches to this dilemma, however, will undoubtedly be tricky to implement.

How Extensive Should an Indicator System Be?

The answer to this question must be determined less by technological than by political considerations—decisions about when the costs of such a system outweigh its benefits. Even though we lack a model that specifies exactly what data we should collect, we have the design, data-collection, and analytical capacity to collect and analyze data about schooling. Within the not-too-distant future, we are likely to have the measurement technology necessary to assess the components of schooling of greatest interest to policymakers and educators.

High costs, substantial reporting burdens

On the other hand, an extensive indicator system would be very costly. For example, it cost about $1.2 million for the most recent surveys administered to a national sample of school principals and teachers, asking them to spend about 45 minutes providing very basic information about school resources, staffing, programs, and student characteristics. The National Assessment of Educational Progress estimates that it currently spends about $4.8 million to obtain a nationally representative assessment of student achievement. To develop a national indicator system, such surveys would need to be augmented by data collection efforts targeted at state- and district-level education officers and perhaps such other efforts as textbook analyses and parent surveys. Moreover, each new school feature (e.g., school discipline practices, homework policies, textbook quality, the use of computers, inservice training) or data source (e.g., parents, textbooks, school board members) that is added magnifies not only the financial costs of the system, but also increases the burden placed on those within the system who are being asked to respond.

The extent of *comparisons* will probably also be limited by cost and burden considerations. Initiating and maintaining a national system that would permit comparing local school systems would be a huge and costly enterprise. Even state-by-state comparisons would increase costs considerably over a system that simply provided nationally

representative information. For example, the National Assessment of Educational Progress recently estimated that altering its current design to permit state-level data would cost an additional $10 million. The costs of increasing national surveys of school principals and teachers to generate data about individual states as well would increase their costs in similar proportions. And with these changed samples and more respondents providing data, the burden on the educational community would increase proportionately. Costs would also multiply with the concomitant increases in the sample sizes needed for each *group* of respondents providing data to the system.

Indicator systems developed within states also entail considerable costs. For example, California anticipates spending about $6.5 million of its approximately $13.5 billion elementary and secondary education budget in 1986-87 to maintain its quite extensive data system, including the development of indicators and performance reports to local schools. Data systems in states with smaller student populations would undoubtedly require fewer resources, and cooperation among states (and between states and the federal government) could reduce costs as well. However, even at the state level, indicator costs are considerable enough to raise serious questions about the cost/benefit tradeoffs of systems of various sizes and scopes.

Timeliness of data critical

Perhaps the most important considerations in the cost/benefit calculation about indicator systems is the value and timeliness of the information that they can provide. Some conditions in the educational system might be long past by the time that policymakers and researchers had organized and analyzed the data resulting from a comprehensive indicator system. More important, as we mentioned earlier, indicator systems should reflect a model of the education system. Without a model that constrains the kind of data collected, it will be difficult for policymakers and analysts to convert data collected *by* an indicator system into information *about* an education system.

Given these considerations, those developing indicators will probably focus attention on a relatively small number of measures that are particularly critical to the well-being of schooling. But because these systems will focus only on *key* indicators, they are not likely to include all of the information that local educators and their school boards, state and national policymakers, or researchers might like. Cost-benefit considerations will probably constrain the comparisons that can be made among states, districts, and individual schools. And finally, tradeoffs will probably be required to balance the comprehensiveness of the indicator system and the frequency with which the data collection is repeated.

What Political Pressures Will an Indicator System Bring?

It should be clear by now that indicator systems will not be neutral, technological information systems, free of political pressures. Choices about what indicators to develop, what purposes indicators should serve, what level of data to collect, and how indicators should be used for comparisons are all political as well as technical issues. But we must also consider the political pressures resulting from the *mere existence* of an indicator system.

Multiple pressures on education community

The most direct pressure from indicators will be felt by members of the educational community. They stand to be judged more publicly than ever before—and with measures over which they have little control. As a result, educators will undoubtedly exert a great deal of pressure themselves to affect the selection of indicators, to influence the level of data aggregation and analysis, and to shape the methods of interpreting, releasing, and reporting data. Once indicators are in place, educators will undoubtedly feel additional pressure to bias responses in ways that might benefit them and their schools—particularly if they feel they have had little say in the development process or that the information will be useless to them. Even now, such pressures have been reported. In some states currently collecting indicator data about school processes, school administrators have been reported to encourage students to "exaggerate" their responses to particular questions about their school experiences. In other areas, informal teacher-networks have spread the word among their colleagues to deflate their salary figures and inflate their teaching-load numbers on state data-collection instruments, reasoning that the resulting data might bring about more favorable policies in both areas These pressures are likely to be proportionate to what can be lost or gained by indicators. If measures are tied to rewards or repudiation, policymakers should expect the pressure to be great.

How will the public respond?

Policymakers should also weigh the effects an indicator system could have on the public, and the pressures these responses might place on policymakers. For example, when indicators make the status and trends in the educational system explicit, how might the information (good or bad) influence public willingness to support schools, generally, or specific educational policy initiatives? More than one scenario is possible. On one hand, the claim is often made that the public will be ready to provide greater financial support for schools—for example, high teachers' salaries—when they are provided evidence that schools and teachers are doing a competent job—raising achievement test scores. But that is not the only possible scenario. If the news from indicators is good, might not the public feel complacent about schools and, as a consequence, make educational initiatives more difficult to mount? On the other hand, the claim is made that an indicator system will sustain public attention and marshal resources for schooling by clearly pointing to components of the system requiring improvement. However, it is also possible, as we have witnessed in the past few years, that w · 'he news about school quality is bad, public support can erode. This erosion might be evidenced both by decreasing willingness to provide resources for education and by increasing numbers of families enrolling children in private schools.

Policymakers might also consider how indicators could influence the discretion they are able to exercise with regard to various educational policies. Once we have good measures and well-publicized data about schooling, pressure for acting on those measures might be substantial. For example, if we developed a "good" indicator of teacher quality and published the resulting data, might not pressures for policies that respond to that indicator result? Such data might increase pressures for merit-based teacher compensation or more "objective" teacher dismissal

processes. Policymakers who favor such policies could have their agendas advanced by such indicator data; policymakers who oppose such policies might find their opposition exceedingly difficult to maintain in the face of "good data." Similarly, provocative data about one feature of the educational system may deflect needed energy and attention away from other components that may be either less easy to measure or less likely to capture public interest. Moreover, if the policy community anticipates public pressure to act on measures in ways that threaten the interests of important constituents, it may well move to suppress the measures or the resulting data.

Indicators and political interests

As we noted earlier, indicator data are unlikely to produce unequivocally good or bad news about schooling. And because most data will be susceptible to various interpretations, pressure will be substantial to interpret and report data in ways that advance the interests of particular groups. Moreover, indicator data are likely to be controlled by those who currently have the most influence in the political system. As a result, indicators have the potential of advancing the political interests of those currently in control and disadvantaging those groups with less control. For example, even though parents are critical schooling constituents, they are li' › have little influence over what data are collected and how t ght be used. As such, indicators can be used to reinforce preexisting power relationships and, perhaps, to actually circumvent significant change.

What we have described here largely reflects "worst case" consequences of indicators. Indicator information, in itself, should not be viewed as dangerous, and most policymakers and education practitioners will undoubtedly attempt to use it wisely and carefully. However, the dilemma of how to explicitly consider potential political pres s in indicator selection, data collection, levels of analysis, and meth ds of disseminating information has yet to be satisfactorily resolved.

Who Should Make Decisions About Indicators?

Definitions of educational indicators follow from views about appropriate educational purposes, the means by which these can best be achieved, and what constitutes proper schooling experiences for children. As we suggested earlier, there is considerable consensus about important school goals and desirable conditions. But indicators must be defined precisely, and these definitions must be operationalized into specific measures. This is where we will undoubtedly encounter difficulty. For example, if the operating definition of student achievement outcomes focused on students' acquisition of basic knowledge and skills in academic subjects, one set of indicators would be appropriate. Scores on traditional paper-and-pencil, multiple-choice achievement measures could serve as indicators of these outcomes; and measures of those educational processes thought to affect these outcomes (e.g., time on task, direct instruction) would emerge as key process indicators. If, alternatively, students' ability to think critically or to apply knowledge in solving real-life problems was specified as the most important achievement outcome, then a different set of key indicators would be appropriate. Nontraditional measures would be needed, perhaps ones that required essay writing or included simulations of problem-solving situations. We

would also want as key indicators, measures of school conditions likely to produce these outcomes (e.g., activity-based lessons, classroom opportunities to solve problems). These two indicator systems could be quite different. Moreover, if we included both types of achievement outcomes, or designated student attitudes, participation in higher education, or parent satisfaction as outcomes to be assessed, resulting indicator systems would vary even more.

Views about these specifics differ among those responsible for the conduct of the educational system. So do opinions about what purposes indicators should serve; what level of information should ground the system; whether or not states and schools can fairly be compared; and whether indicators will be of real value to their professional and public consumers.

Given these significant differences of opinion, who should make the decisions about educational indicators? Should the priorities of policymakers or the judgments of professional educators prevail? Should national goals be preeminent? Should state-level interests? Should local conditions be studied to determine what indicators are most relevant to existing practice? Should practicing educators be surveyed to determine what the outcomes of education should be? Should academics? Should state and local political and business leaders be consulted as to their views of desirable indicator systems? If these various sources should conflict, whose judgments are to be considered most legitimate as the basis for indicators? Should indicators be restricted to those dimensions about which there is consensus among these often competing interests?

These questions are inseparable from the technical side of developing indicators, since decisions about what specific features of the educational system should be measured with indicators depends largely on what we want our school system to achieve.

STATE OF THE INDICATOR ART

We have reviewed the basic types and uses of indicators and suggested some of the political issues that must be considered in designing and using indicators. In this section, we turn our attention to the current scene. We will review the indicators now in use, describe efforts to improve them, and discuss some of the questions entailed in designing a national indicator system.

WHAT KIND OF DATA ARE BEING COLLECTED NOW?

We are not without educational indicators. Currently, a substantial amount of educational data are collected regularly. The federal Department of Education publishes each year *The Digest of Educational Statistics*, a compendium of statistics about education, and *The Condition of Education*, a collection of charts and graphs based on these statistics and augmented with some discussion of their meaning. These publications include national statistics on school enrollments, staffing, and financing; characteristics of public and private elementary schools and teachers; data on institutions of higher education; and information about where students go after high school. Some statistics are broken down by regions of the country and sectors of the population; some trends are reported. Neither publication, however, attempts to present a system of indicators. Moreover, the quality of the data collection efforts supporting these publications has been severely criticized.

Studies sponsored by the Department of Education (e.g., the ongoing National Assessment of Educational Progress, the recent High School and Beyond longitudinal study, and the upcoming National Educational Longitudinal Study of 1988) gather and report either cross-sectional or longitudinal information about student achievement and educational conditions in a representative sample of schools. These data can be used to draw conclusions about schooling nationally. However, they have not been designed to feed information regularly into an indicator system or to provide state-by-state data.

In addition to these federal efforts, most states collect a great deal of information about school finance, enrollment, and achievement. Some states have developed indicator systems for evaluating and comparing local districts. California, for example, measures and reports "quality" indicators including student achievement, dropout rates, student participation in Advanced Placement classes, and the number of writing assignments students are given.

As we noted earlier, the closest existing approximation of a national indicator system is the Secretary of Education's controversial "Wall Chart" consisting of a small number of statistics drawn from federal data collection efforts or aggregated from state data.

IMPROVING INDICATORS

Researchers and practitioners at every level are working energetically to develop an improved set of national educational indicators. Efforts are currently under way at the federal Department of Education (through its Center for Statistics), the National Governors' Association, the Council of Chief State School Officers (through its Assessment Center), and the

National Science Foundation (through its sponsorship of projects at such research centers as the National Academy of Sciences, The RAND Corporation, and the University of Wisconsin). Indicator development is also going on within many state education agencies.

Critical need for better measures

Perhaps the most obvious challenge to these efforts stems from the paucity of adequate educational measures. For example, most would agree that the quality of the teaching force should be a key indicator. But, while we can easily measure teacher *qualifications* (such as credentials, training, and years of experience in the classroom), these don't really capture teachers' ability to teach. However, we lack direct measures of teachers' skill in the classroom. A second example of a component with inadequate measures is student achievement, the component many consider the most important for indicators. Here we have fairly good paper-and-pencil measures of the most commonly taught basic knowledge and skills. But we lack adequate measures of children's ability to think critically, to apply their knowledge, or to solve problems. A third area of needed development is curriculum. Of all the components of the educational system, we probably know least about measuring the curriculum that students experience at school. Before an indicator system can be useful in assessing these and other significant components of the system, new measures must be developed.

Consistency also an issue

In addition to pointing to areas where new indicator development is needed, indicator projects are attempting to identify existing measures that meet the "good-indicator" criteria listed earlier. Efforts are also under way to arrive at common definitions of important factors--for example, what is a dropout? who is an attender? Deliberations among the states aim at producing consistent methods of collecting data. Because states now use quite different measures, results cannot be combined into national indicator data.

Some projects are also considering essential questions related to the design and implementation of an indicator system: What level of data should be collected about various schooling components? How might ongoing federal and state data collection efforts be modified to meet the needs of an indicator system? These are complex and difficult conceptual and technical matters. They are not likely to be easily or quickly resolved.

WHAT TYPE OF INDICATOR SYSTEM IS LIKELY TO EMERGE?

There is no generally agreed-upon master plan for the development of indicators. However, efforts at the federal and state levels appear to be directed toward the development of a "mixed" two-level system of educational indicators, such as that shown in Fig. 19.

Two-level system most likely

The first level of the system will probably consist of national indicators measuring schooling components across states and localities. These indicators will permit us to observe and monitor nationally *core* features of schooling. Some national-level indicators are likely to consist of data that are collected regularly by the federal government itself, probably through existing data collection efforts like the surveys administered by

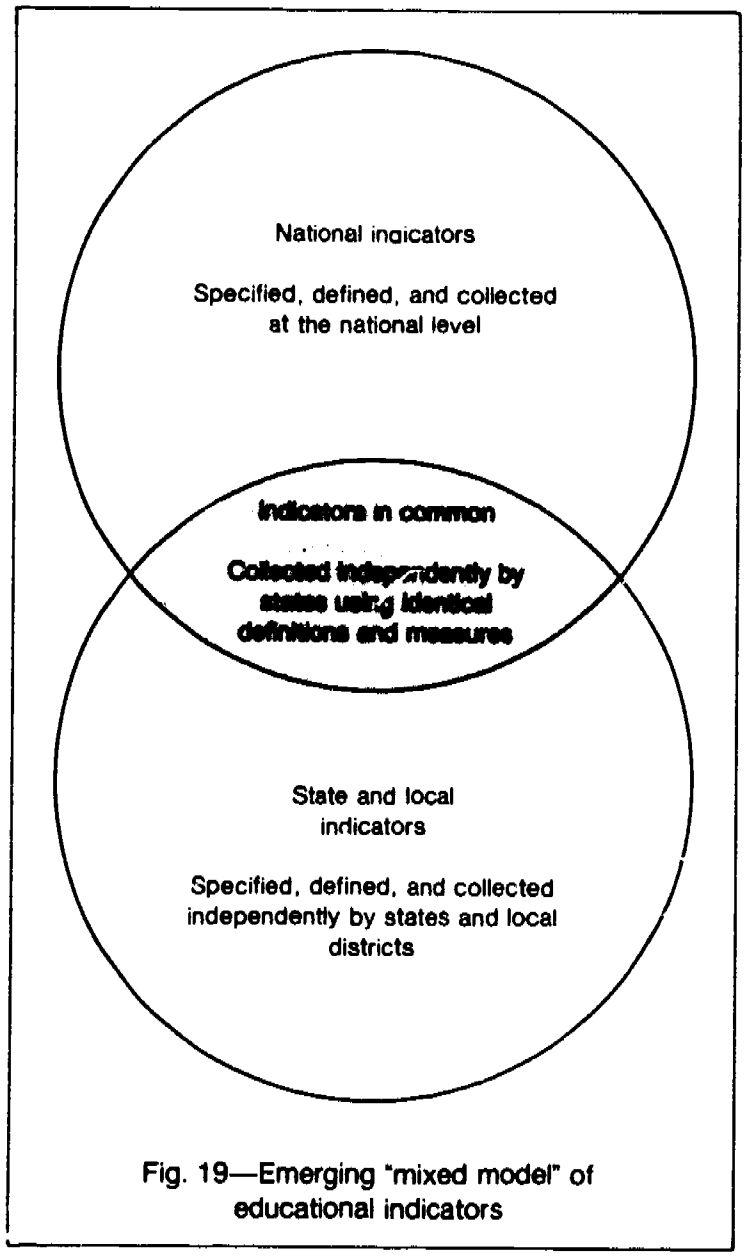


Fig. 19—Emerging "mixed model" of educational indicators

the Bureau of the Census, the National Assessment of Educational Progress, and National Surveys of Public and Private Schools and Teachers. Some additional indicators of conditions particularly relevant to mathematics and science education may be generated by data collection efforts sponsored by the National Science Foundation. Some of these federally generated data will probably be collected so that they can be broken down on a state-by-state basis. Other indicators will undoubtedly represent a mix of national and state-level indicators since they will be formed from data collected by the states and based on agreed-upon definitions of indicators and common measures.

The second level of the indicator system will probably be composed of separate indicator systems within states. These systems will be tailored to the goals and needs of each state and used to supplement those data collected for the national indicator system. In some states, these state-

level indicator systems will include provisions for locally developed indicators to meet the information needs of local schools and districts.

WHEN MIGHT AN INDICATOR SYSTEM BE IN PLACE?

Even if we currently had what we thought was an adequate set of educational indicators and a plan for designing and implementing a system, the evolution of a *mature indicator system* will need to take place over time. This will be an interactive process. Tentative indicators must be used and their results studied over time to evaluate their worth and to develop better indicators. Until various indicators are tried out, modified, and retried, we will not know which ones will prove valid and most useful. As a result, a complete and sophisticated indicator system may take decades to develop.

However, work currently under way is likely to produce an *improved set of indicators* in a much shorter period of time. Through these efforts, the existing "Wall Chart" statistics could soon be obsolete.

MATCHING THE PROMISE AND THE REALITY

In an era dubbed "the information age," we should not be surprised that many believe good statistical indicators will provide the information necessary to understand and improve our schools. That confidence rests on our ability to identify and analyze key schooling elements; and our skill in selecting as indicators those features that are fundamentally stable and predictable, and, therefore, appropriately subjected to repeated assessment. Our hopes for indicators are also based on our ability to quantify and describe schooling features free of bias, values, and opinions, and to measure, report, and interpret data separate from the social and political struggles among various interest groups. If we can meet all of these requirements, we expect that better control, quality assurance, and efficiency can be obtained with scientifically designed indicators. We optimistically expect that more rational processes will guide educational policymaking, and that judgments about educational quality will be less subject to political trends and temperaments.

Much of this confidence, and our hopes for indicators, may be exaggerated. Educational organizations may be far more complex, dynamic, and interacting than data about discrete, or even mechanically-linked parts, can convey. We may have a great deal of trouble capturing and measuring their most important features. "Successes" detected by indicators may be so short-lived that, by the time they are reported, they no longer exist. Educational planning based on indicator data may fail to bring about intended results and have contrary, unintended, and unpredictable consequences. Moreover, it is hard to imagine that collecting, reporting, and interpreting indicator data will be free of political influence.

The intent here is not to argue against indicators. Rather, it is to suggest that, even as we pursue indicators, we must acknowledge their limitations. As with all human inventions, indicators' contributions to schooling will depend on the thoughtfulness with which they are designed and applied to problems. Indicators can be no more than tools to aid the ongoing political dialogue about what we want from our schools and how those ends can best be accomplished. They will not carry with them a single interpretation of past events, offer clear judgments about

the present, or provide direct answers about what should be done next. They will, at best, bring new knowledge to bear on educational issues, stimulate more thorough discussion and debate, and suggest creative new solutions to problems. But if indicators are allowed to serve these more modest purposes, their contribution is likely to be substantial.

INFORMATION SOURCES

The following people and organizations are currently involved in indicator development. They may be contacted for further information.

For information about efforts to develop common state measures and link data collected by states:

Ramsay Selden, Director
Center for Assessment and Evaluation
Council for Chief State School Officers
377 Hall of the States
Suite 379
400 North Capitol Street
Washington, D.C. 20001

Evelyn Ganzglass
Center for Policy Research
National Governors' Association
Hall of the States
444 North Capitol Street
Washington, D.C. 20001

For information about national indicator systems:

Emerson Elliott, Director
U. S. Department of Education
Center for Statistics
555 New Jersey Avenue
Washington, D.C. 20208

Marshall Smith, Dean
Graduate School of Education
Stanford University
Stanford, California 94305

Susan Fuhrman, Director
Center for Policy Research in Education
Rutgers University
Eagleton Institute of Politics
New Brunswick, New Jersey 08901

For information about state-level indicators:

Alex Law, Director
Division of Program Evaluation and Research
California State Department of Education
721 Capitol Mall
4th Floor
Sacramento, California 95814

Kevin Crowe, Director
Planning, Research, and Evaluation
Nevada Department of Education
400 W. Capitol Complex
Carson City, Nevada 89710

J. Robert Coldiron
Ross S. Blust
Division of Educational Testing and Evaluation
Pennsylvania Department of Education
333 Market Street
Harrisburg, Pennsylvania 17126

For information about mathematics and science indicators:

Thomas Romberg
Department of Curriculum and Instruction
University of Wisconsin-Madison
Madison, Wisconsin 53706

Senta Raizen
Commission on Behavioral and Social Sciences and Education
National Academy of Sciences
2101 Constitution Ave., N.W.
Washington, D.C. 20418

Richard Shavelson
The RAND Corporation
1700 Main Street
P.O. Box 2138
Santa Monica, California 90406-2138

This paper may be obtained directly from the Publications Department, The RAND Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138. Telephone (213) 393-0411, Extension 6686.